THE X-RAY ANALYSIS OF COMPLEX CRYSTAL STRUCTURES: α -RHAMNOSE MONOHYDRATE

Thesis presented to the University of Edinburgh
for the degree of Ph.D.

by

H. McD. McGEACHIN

May, 1956

ABSTRACT OF THESIS

Name of Candidate Hugh McDougall McGeachin

Degree Ph.D. *Date* 14 th May 1956.

Title of Thesis The X-ray analysis of complex crystal structures: α- Rhamnose Monohydrate.

The crystal structure of α-rhamnose monohydrate has been determined by the interpretation of the three dimensional Patterson function, and refined on two projections using diffraction data derived from Geiger counter measurements. There are possibly significant deviations from the expected values of three bond lengths, which include both the axial C-OH bonds of the molecule, one of which is short, as in α-glucose, and the other long. There are several significant deviations from the tetrahedral angle in the bond angles; that at the oxygen atom of the ring is 120° and those made by the extracyclic substituents appear to indicate a repulsion between them. The crystal packing is determined by hydrogen bonds, in which all the oxygen atoms and the water molecule take part, and in which all the available hydrogen atoms are utilised. The hydrogen atoms appear with moderate definition on $(F_o - F_c)$ syntheses but their positions are not completely unambiguous, largely owing to the non-centrosymmetrical nature of the crystal, the poor resolution of the atoms in projection and the occurrence of anisotropic thermal vibrations of the atoms, for which no corrections were made.

Use other side if necessary.

THE UNIVERSITY *of* EDINBURGH

The field of structural carbohydrate crystallography, although strangely neglected, is of considerable interest and importance in several branches of scientific research, chemical and biochemical. While the molecular formulae of the monosaccharides have been elucidated with reasonable certainty by purely chemical methods, there remain ambiguities about even the gross features of the spatial distribution of the atoms (Reeves 1950, 1951). This is a consequence of the puckered nature of the six-membered pyranose ring which, it is supposed, is the structural basis of all the monosaccharides; for by changing all axial substituents to the equatorial positions, and vice versa, the cis-trans relationships, which are all that are indicated by most chemical experiments, are preserved, but the molecule has, nevertheless, a completely different shape. This fact, which appears to have been generally unremarked for many years, led Reeves to examine the powers of different sugars to form cuprammonium complexes; analysis of his results enabled him to formulate a series of rules by which the relative stabilities of the different conformations could be predicted. Similar complex-forming experiments by Boeseken (1949) led to similar conclusions.

Reeves' rules, nevertheless, have been greeted with considerable suspicion by carbohydrate chemists, partly, no doubt, due to their purely empirical nature. Attempts have been made (Schwartz, unpublished) to give a better theoretical basis to this conformation analysis by assuming that the interaction of the dipole moments of the carbon-oxygen bonds can account for the relative stabilities of the different conformations. A detailed study of the electron density distribution in these bonds, such as could be achieved by X-ray analysis pushed to the limits of accuracy possible, would obviously be of the highest value in examining this hypothesis.

Chemically, the sugars possess some puzzling features. While the evidence for their ring structures is overwhelming it is generally supposed that in solution there is some of the open chain form present, which could account for the reducing properties of aldehydic sugars. However since this reducing power is greater than that which would be expected even if all the molecules were in the straight chain form (and also since ketonic sugars are also vigorous reducing agents),

another explanation of this effect must be sought. MacDonald (1950) has suggested that this phenomenon may be connected with the abnormally short C - O bonds which he found in the X-ray analysis of α-glucose. Before such a speculation can be promoted to a theory it is obviously essential to find out whether the occurrence of such short bonds is a general feature of sugars or merely a local peculiarity of the glucose structure. Another point of interest, to investigate which a detailed knowledge of the electron distribution is required, is the explanation of the unequal reactivities of the hydroxyl groups; it is not known at present whether this is an electronic effect or whether it is merely due to steric availability.

The phenomenon of fermentation is one in which an accurate knowledge of sugar structures might prove of value. The action of enzymes and bacteria is highly specific to structure, so much so that this constitutes one of the principal methods for their identification. While the mechanism of fermentation processes is almost certainly complex and difficult to elucidate, the structure of the sugar involved

plays such an important part that a detailed knowledge of it would be an essential preliminary to any convincing explanation.

Apart from possibly shedding light on questions of the mechanisms of chemical reactions, the accurate determination of the structures of a series of sugars would produce valuable data on bond lengths and bond angles. The fashionable present-day practice of proposing structures for macromolecules such as proteins, viruses and polysaccharides by building models conforming to general principles such as the maximum degree of hydrogen-bonding and taking advantage of all the scanty experimental evidence, depends very largely on these fundamental data; the whole approach to the problem would be very much strengthened if the permissible deviations from the standard values were better defined. This applies in particular to bond angles; not only are they likely to vary more than bond lengths, but small variations in them are likely to cause a relatively much greater effect in the structure of a large molecule. It should be noted that deviations from the tetrahedral angle in a polysaccharide should be of the same order as those in a monosaccharide, since they

are both presumably caused by the strain due to packing by hydrogen bonding. A knowledge of the angles actually occurring in a series of sugars, or in different crystalline modifications of the same sugar, would thus be of great value in proposing such structures, and might help to serve as a criterion for discriminating between rival proposals.

In spite of these pressing reasons for the complete structural determination of sugars, very few complete analyses have been carried out. This is all the more surprising since their importance was realised in the early days of organic crystallography. Cox and his co-workers collected unit cell and space group data on a large number of free and methylated sugars (Cox 1931, 1932; Cox et al. 1932, 1935, 1936). Assuming that in certain cases there exists a correlation between the sizes and shapes of the molecule and the unit cell, particularly in the cases of the methylated sugars where hydrogen bonding cannot be the predominating factor in the molecular packing in the crystal, they drew certain conclusions from the frequent occurrence of a short axial length of 4.5 - 5 A^{o}, which they took to

correspond to the thickness of the molecule. Now if the strainless Sachse-~~Mo~~Mohr pyranose ring is the structural basis of monosaccharides, a constant molecular thickness throughout the sugars would not be expected. β -glucose, for instance, would be thin, since all the hydroxyl groups are equatorial; α -glucose, having one axial group would be thicker and α -mannose, having one axial group on each side of the ring would be thicker still. The differences observed in the short axial lengths of thick and thin pyranoses were much less than would be expected if the ring is a Sachse-Mohr type. Cox therefore suggested that the ring was not of the classical type based on tetrahedral angles, but consisted of a nearly planar group of five carbon atoms, the ring being closed by the oxygen atom which is not in the plane. If small deviations from tetrahedral angles are allowed this ring is strainless. A molecule based on this ring will not be as thick as one based on a true pyranose ring, and the difference in thickness between thick and thin sugars is much less (0.26 A as against 1.0 A). This difference was suggestively close to those observed in unit cell dimensions and this

constituted very convincing evidence for Cox's theory. Nevertheless, all subsequent crystal analyses show that the ring actually occurring in sugars is of the classical Sachse form.

The following structural analyses of sugars have been published:-

Glucosamine hydrobromide (Cox and Jeffrey 1939).

Sucrose sodium bromide (Beevers and Cochran 1947).

Glucose (MacDonald and Beevers 1950, 1952; MacDonald 1950).

Sucrose (Beevers et al. 1952).

Difructose strontium chloride trihydrate (Eáland and Pepinsky 1950).

In addition the structure of cytidene (Furberg 1950) which contains a ribofuranoside residue has been determined.

Of these, the glucosamine and fructose compounds have only been solved on one projection and cannot be used as sources for bond lengths and angles; the sucrose sodium bromide structure was not refined to any great accuracy and although of great importance as the first sugar structure to have been completely determined by X-ray analysis, and indeed as the only perfectly unequivocal proof of the chemical

structure of sucrose, can probably not be relied upon to give accurate dimensions (although it seems to have been preferred to cytidene to provide data on the furanose ring for model-building experiments on nucleic acids). The structure of free sucrose is still under examination. This leaves glucose as the only pyranose sugar to have been studied with reasonable accuracy until the present work on rhamnose. A comparison will be found later.

The reasons why so few crystal structures of the sugars have been determined are not fully understood by the candidate. The crystals, although often difficult to grow, usually give excellent diffraction patterns to the limit of Cu $K\alpha$ radiation. The structures are admittedly not easy to solve; the lack of symmetry and contorted shapes of the molecules, and the low scattering powers of the atoms involved usually result in not very distinctive diffraction patterns which make trial and error and Fourier transform methods difficult and phase inequality methods impossible, while the solution of the Patterson function is complicated by the occurrence of equal and parallel vectors. The crystals, too, unless DL forms can be

prepared, are necessarily non-centrosymmetric. This, in general, however, does not make the solution of the structure more difficult (unless Fourier transform methods or direct solution by inequalities is the technique employed) and indeed in the solution of the Patterson function a rotation or screw axis, or a mirror or glide plane is always a preferable symmetry element to a centre; it adds complications to the refinement, however, and decreases the accuracy of the structure. In spite of all this, monosaccharide structures can, with patience, be solved and it seems a pity that the investigation of more of them is not undertaken.

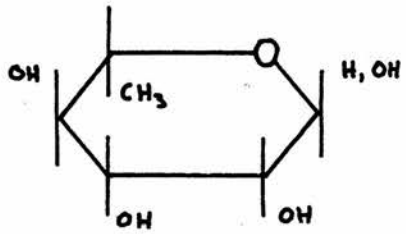
α-Rhamnose

Rhamnose is 6-desoxy mannose and so is represented in the open chain form by

$$CH_3 - \overset{OH}{\underset{H}{C}} - \overset{OH}{\underset{H}{C}} - \overset{OH}{\underset{H}{C}} - \overset{OH}{\underset{H}{C}} - CHO$$

Hirst and Macbeth (1926) by oxidising trimethyl rhamnose to arabotrimethyloxyglutaric acid whose structure was known, showed that the methylated sugar certainly, and hence the free sugar probably, had the six membered pyranose ring. At this time this was rather uncertain and indeed a furanose form had been proposed (Hudson, 1926) because of its anomalous rotatory power.

The probable cyclical formula can thus be assigned:



There is considerable evidence for the position of the hydroxyl group on carbon atom 1 in the α form. The complex-forming experiments of Boeseken with boric acid mentioned above indicate

that the hydroxyl groups on carbon atoms 1 and 2 are trans, and this evidence is supported in a general way by the study of optical rotations; in L sugars the α form has the lowest positive rotation and this is generally supposed to be possessed by the configurations which have the substituents on carbon atoms 1 and 5 trans to each other, in the cyclic formulae. In the case of rhamnose however, since its optical rotation is so peculiar, such evidence cannot be regarded as conclusive. Reeves(1951) rules also support the trans configuration.

α-Rhamnose crystallises beautifully as a monohydrate, by evaporation of a solution of methyl alcohol (90%) and water (10%). The anhydrous sugar has also been crystallised (Charalambous 1953) with some difficulty.

Experimental

The crystals were examined by an optical goniometer and the crystallographic axes given by Groth (1910) identified. It may be mentioned that the predominating crystal habit was different from all three illustrated by Groth. Specimens for X-ray work were prepared by cutting prisms about the required axes (with a razor blade or with moistened thread) and rounding off the edges by further cutting, rubbing with moistened filter paper and rolling on a ground glass plate. By these means reasonably cylindrical specimens could be produced without too much trouble.

Oscillation photographs about the different axes were taken to give approximate values of the lattice translations and to adjust the crystals for subsequent photography. Weissenberg photographs of layer lines up to the fourth were then taken about the three axes using a normal beam camera of 10 cm. diameter (Beevers 1952). These were readily indexed by inspection; it is the candidate's view that the difficulties of indexing normal-beam Weissenberg films of upper layer lines have been highly exaggerated.

Accurate cell dimensions were determined by the extrapolation technique of Farquhar and Lipson (1946). Measurements of the distance of high order reflexions from the shadow of the knife edge of the camera were converted to the Bragg angle θ by means of a graph obtained by experiment with a quartz powder sample. Reflexions of the type $h00$ were used first to give values of $\underline{a}^*$ $\underline{b}^*$ and $\underline{c}^*$. These values were then used for reflexions of the type hko; when for instance h is small, k will be large and so the value of S^2 ($= h^2 \underline{a}^{*2} + k^2 \underline{b}^{*2}$) will depend largely on $\underline{b}$. The values of b so obtained were plotted against $\sin^2\theta$ and the points used for an extrapolation to $\sin^2\theta = 1$. When $\underline{a}$ and $\underline{c}$ had been determined a similar method was used to find $\cos\beta^*$. Since S is very sensitive to changes in cell dimensions in high order reflexions this procedure usually gives accurate results: the extrapolation minimises errors due to possible film shrinkage. The results obtained were, taking the wavelengths of Cu Kα radiation as $(K\alpha_1) = 1.5405$ A

$(K\alpha_2) = 1.5443$ A

$\underline{a} = 7.910$ A $\pm .001$

$\underline{b} = 7.914$ A $\pm .005$

$\underline{c} = 6.674$ A $\pm .004$

$\beta = 95^\circ 37' \pm 5'$

These are in excellent agreement with the figures published by Cox (1931) and Andress and Reinhardt (1931).

	Cox	Andress & Reinhardt
a	7.98	7.86
b	7.97	7.86
c	6.73	6.63
β	95°25'	95°

Cox's values of β is taken from Groth: both these sets of cell dimensions are derived from layer line measurements. These figures have been converted from kX units to A°.

The density was found, by floating small crystal particles in a mixture of benzene and chloroform, to be $\rho = 1.457 \pm 0.005$. This gives the number of units of $C_6H_{12}O_5\ H_2O$ in the cell as 2.005 = 2.

The symmetry of the reflexion intensities indicated that the crystal was monoclinic; the only systematic absences observed were oko for k odd and so the probable space group is $P2_1$.

Intensities of the hko, hol, and okl zones were carefully estimated by comparison with low order reflexions on standards which had been exposed for different times under as near as possible the same conditions. The same crystals were used for the standards as for the original

exposures so as to reduce errors due to differences in spot shapes between crystals, at least for the low orders, where, because of the great density of the spot, accurate estimation is difficult. By the use of multiple exposures, all observed intensities could be brought into the range where accurate comparison is possible. Several different crystals were used for these zero layer line photographs, all approximately cylindrical with diameter of 0.1 - 0.2 mm. Since in the (hol) zone however, there were several reflexions unobserved (because of the effective centre of symmetry in the b -axis projection) an exposure was made with a much larger cylinder (about 0.6 mm. diameter) which reduced the number of unobserved reflexions considerably. In general more accurate intensities can be obtained from small crystals, particularly when it is desired to ignore absorption corrections, but a greater fractional error in very small intensities is admissible.

The intensities were not corrected for absorption; calculation indicated that the effects of this would be of the order of 1%, which is much less than the accepted values of experimental error in visual estimation, and so could be ignored.

The intensities were corrected by Lorentz and polarisation factors using the graphical method of Cochran (1948), and placed on an approximately absolute scale by Wilson's (1942) method.

When it became obvious that the structure was not going to yield to a two-dimensional attack, the general (hkl) intensities were estimated by a much faster but less accurate method. This depends on the fact that if the hko, hol and okl reflexions are known, these provide standards on any upper layer line Weissenberg photograph: on the h1l photograph, for instance, there occur the reflexions h1o and 01l which are known. The corrected values for these reflexions were used to estimate the F^2_{hkl} directly of spots which had approximately the same value of θ. The values obtained by this method will not be accurate but it was thought (with justification) that they would be sufficiently good for coefficients for a three dimensional Patterson function. Because of the lack of computing facilities, three dimensional refinement was not contemplated.

At quite a late stage in the work on rhamnose the laboratory was fortunate enough to

obtain a Geiger counter with associated electronic circuitry. The counter was mounted on a diffractometer constructed (largely, with great kindness, by Messrs Ferranti Ltd.) from the prototype of the Beevers normal beam Weissenberg camera. This allowed the crystal and counter to be rotated independently and to be set to any desired position with considerable accuracy. The geometry of the X-ray set unfortunately set a limit to the Bragg angle which could be investigated as is usual in such experimental arrangements.

The counter fed pulses to a quenching preamplifier (Type 104A) mounted on the same bracket which in its turn fed them to a scaler (Ekco N 530) through a sensitive galvanometer. This served as a ratemeter by which the reflexions could be readily located; the actual measurements were made on the scaler, which had 'auto-time' and 'auto-count' circuits and also provided the H.T. supply for the counter.

The X-ray set used was a pre-war Metropolitan-Vickers demountable set, which, while by no means the ideal source of X-rays for work of this type, was known to be considerably better behaved than the more modern Raymax tubes.

A constant voltage transformer was fitted to the filament supply circuit and the set was rectified in order to keep the effective dead-time of the Geiger counter as low as possible (Cochran 1950). From a series of observations made over a length of time on one reflexion from a crystal it was found that on the whole the set was satisfactory; a very necessary precaution when making intensity measurements on crystals however was to check the output of the set before and after each measurement. This took account of two effects - the contamination of the target, which cuts down the X-ray output to a much more serious extent than had hitherto been realised in this laboratory (by as much as 20% in a week with a newly cleaned target and a current of 10 m A), and an occasional sudden discontinuity in the X-ray output which was attributed to a sudden change in the position of the focal spot, caused by some event in the filament. When this took place during the measurement of a reflexion, the reading was discarded and a re-measurement made. A curious phenomenon showed by the tube was that while running at 50 k V the X-ray output was proportional to the current passing through the tube, as is expected, at 30 k V (the voltage used, so as to reduce the white radiation) it increased with current up to 10 m A, flattened out

and then decreased with increasing tube current. This effect is presumably due to a change in size and shape of the focal spot, which the focussing arrangement of the tube is unable to correct at this reduced voltage: but the matter is not properly understood.

The dead time τ of the counter was found by taking readings of the counting rate with a variable number of nickel foils in front of the window. These foils were all cut from a large sheet to ensure, as far as possible, uniformity in their thickness and were tested, using the counter, for a uniform absorbing power. Their thickness was measured with a micrometer; the mean of 20 readings taken at different places and on different foils was .00065" which, however, is probably not very accurate; this gives the factor m by which Cu kα radiation is cut down as 2.05, the error in which, assuming the error in measurement of the thickness as 10%, is $\pm$ 0.15. If the incident X-ray beam is constant throughout these readings, and if it is sufficiently monochromatic for the different absorption factors for white radiation to be neglected, then

$$\frac{N_0}{1-N_0\tau} = \frac{mN_1}{1-N_1\tau} = \cdots = \frac{m^nN_n}{1-N_n\tau}$$

where Nn is the observed counting rate with n foils in position. Up to five foils were used, and taking the calculated value of m as 2.05 as the starting point of an iterative process, final values of m = 2.07 and τ = 352 ± 15 μsecs were obtained.

Tests were then made to examine what technique could be used to obtain the integrated intensities of the reflexions. Cochran (1950) has described a method which he found to give excellent results which consists of rotating the crystal through a degree or so while counting; he finds that by multiplying the dead time τ by a factor k, which he took constant over all values of θ, he can make allowance for the different dead time corrections which are applicable to the different counting rates found as the crystal is rotated through its reflecting position. As he points out, this factor is dependent on the shape of the spot, but he finds that no great error is introduced by assuming it independent of θ, *i.e.* independent of spot shape, providing that the counting rate is not too high. Experiments were carried out on this matter by measuring the total number of counts registered on the scaler in time T while sweeping the crystal, driven slowly by an electric

motor, through its reflecting position with a variable number of foils in position. The value of $\frac{K_1 t}{T}$, the effective dead time, could thus be deduced for reflexions with different values of θ ; it was found that the values obtained were by no means constant.

θ^o	$K_1 t/T$ μsec.
21	44
27	67
27	60
28	68

The last three values show that this is not an effect due to irregular crystal shape since the measurements were made at different angular positions of the crystal in order to test this, and the only conclusion that could be reached was that in the experimental arrangement used here, the sweeping technique could not be used to give intensities of the highest accuracy.

It was felt, with reluctance, that in the circumstances the only appropriate technique was the immensely laborious one of taking readings of the counting rate at sufficiently close intervals from a stationary crystal, correcting these separately, plotting them against the angle and counting the number of squares under the

resulting curves. The apparatus was adjusted so that the focus was fairly fine, so that the width of the profile was about $\frac{1}{2}^{o}$; measurements of the background were made on both sides of the reflecting position so that the area under the curve due to the characteristic radiation could be well defined. The background on the whole was low, ~~usually~~, when two or three foils were in position, which was always the case unless a weak reflexion was being measured, of the order of 5 counts/sec. The observed counting rate was not allowed to rise above about 500 counts/sec, as then the accuracy of the dead time correction is impaired. The slit in front of the counter window which was 4 mm. wide, enough to receive a complete reflexion of the highest orders which could be reached, was mounted eccentrically. This was because it had been found in previous experiments with a narrow (< 1 m.m.) slit that the central wire of the counter absorbed X-rays, which were thereby lost. This mounting also increased the upper limit of θ attainable by a few degrees.

Measurements of the intensities of the hol and hko reflexions were made by this technique; particular care was taken to correct the readings to the same scale by measurement of the maximum of a standard

reflexion before and after, as discussed earlier. A series of measurements on one reflexion, taken at intervals throughout the work indicated that the standard deviation of a single measurement, which includes errors due to correction, plotting and area measurement, was of the order of $2\frac{1}{2}\%$.

Care was also exercised in the correction of the intensities and the graphical method of obtaining $1/\Theta$ was replaced by the use of tables for the low order reflexions where the % accuracy of the graphical method is low.

The Solution of the Structure

A short account of the solution of the structure has been published in Spanish (McGeachin 1954). A variety of methods were tried and an outline will be given here of the history of the work, with details only when the methods are not those in most common use.

The main attack on the structure was by Patterson methods and initially the projections down the three axes were computed on an analogue computor (Stern 1950) of the Hagg-Laurent (1946) type. These were used to try to find the orientation of the molecule. The pseudo-symmetry of the pyranose ring causes its vector set to have a very simple form (Fig. I), consisting of six coplanar vectors at angles of 60° of 2.5 A length with six non coplanar vectors of 1.5 A length bisecting these angles and inclined at alternately plus and minus $19\frac{1}{2}^{\circ}$ to the plane. The 2.5 A vectors were easily identified on the projections and showed that the 'plane' of the molecule was tilted to all three axes. (By the 'plane' of the molecule is meant one of the two parallel planes on which three of the six ring atoms lie.) The 1.5 A vectors were not resolved but of the two possible orientations of these vectors, one was

preferable to the other. When the plane of the ring has been determined by the 2.5 A vectors there are twelve possible orientations of the molecule - two sets of six related by rotations of 60°, one corresponding to each way up of the molecule. The sets are not equivalent in a monoclinic crystal unless the plane of the molecule is perpendicular to the b axis. When the puckering of the ring has been determined by the 1.5 A vectors, each set is reduced to three, related by rotations of 120°. Using the preferred orientation of the 1.5 A vectors, it was found that of the six orientations of the molecule, four would not give acceptable packing along the short c axis, assuming reasonable values for the minimum distances of non-bonded atoms (> 3 A unless there is the possibility of hydrogen bonding) and assuming of course that the model constructed on chemical evidence was correct.

Attempts were made to decide between these orientations by examination of the Patterson projections but as their vector sets differed in only four vectors, it was not surprising that the choice proved impossible. Attempts to find the position of the ring in the cell by the heavy vector (at least sixfold) which gives

the distance between the centres of the rings in the centro-symmetrical b axis projection, were likewise unsuccessful. These were supplemented by trial and error methods in which the F_{hoo} and F_{ool} were calculated as functions of X_1 Z_1 the coordinates of an arbitrary origin in the molecule; these proved inconclusive, however, although one orientation (here called A) gave somewhat better agreement than the other (B). The potential contribution of the water molecule was a source of considerable uncertainty in this work. It was later noted when the structure had been solved that extinction of some of the reflexions was the primary cause of failure of these simple methods; even at the time it was realised that dangerously few experimental data were being utilised, although the large extinction errors were not anticipated. The F_{oko} which do not depend on the position of the molecule in the cell were also calculated for both orientations; again A was preferred.

Inequality methods were tried on the F_{hol}; they resulted in the probable sign for one reflexion and were abandoned.

The complete three-dimensional Patterson function was then computed. The 2.5 A vectors

were clearly and the 1.5 A vectors partially resolved, thus now fixing the puckering with certainty. Parts of special sections through the function are shown in Figs. II and III, showing these vectors. The four vectors which differ in the two orientations were examined and again A was the more satisfactory. Attempts were made to find the location of the molecule in the cell by interpretation of the Harker section $y = \frac{1}{2}$; the Harker vectors of the type $2x, \frac{1}{2}, 2z$ which occur between atoms related by the screw axis are only half the weight of vectors between unrelated atoms and this, as is customary, made the Harker peaks insignificant enough to prevent their location by trial and error fitting of a projection of the molecule drawn to twice the scale. The Patterson section at $y = {}^{28}/_{60}$ was subtracted from the Harker section in the hope that peaks due to vectors between unrelated atoms which do not necessarily occur at exactly $y = \frac{1}{2}$ might be eliminated, giving a better chance of identifying the Harker vectors, but this too was unsuccessful.

Patterson superposition methods were then adopted. The relative positions of the six atoms of the pyranose ring were known with some certainty and these can be used as

searchers in making superpositions, the problem then being to look for the extra-cyclic atoms and the other molecule. The screw axis can then be placed half way between the molecules and a search can be made for the water molecule. Because of the centro-symmetric pattern of the searcher atoms, however, a spurious centre of symmetry is introduced into the superposition maps at the centre of the ring and since a structure like a sugar which contains many equal and parallel vectors usually introduces a large number of spurious possible atomic sites, the interpretation of the maps resulting from the sixfold superposition was difficult. In order to reduce the number of possible sites, only the parts of the Patterson function whose value was greater than zero when F^2_{ooo} was omitted were used and in such circumstances it may happen that a true atomic site will not be revealed if the method is used blindly; this actually occurred to a small extent since some of the carbon Harker vectors lay below this level. The great advantage of a screw axis in the interpretation of superposition maps is the certain knowledge that two related atoms lie at levels differing by $y = \frac{1}{2}$. It

is thus sometimes possible, by comparing all possible sites on sections of the maps at y and $y+\frac{1}{2}$ to obtain quite a small set of positions for the screw axis one of which, the true one, should occur in all sets. If two atoms occur at the same value of y, a very powerful restriction is placed on the position of the screw, since the vector between the two sites must be the same at both levels. This technique was adopted but because of the very large number of spurious peaks and the non-occurrence of others, no conclusion could be reached from the superpositions of the ring atoms alone. As complete confidence was not held that the orientation A was certainly correct (indeed, at this stage, some doubts were felt as to the actual structure of the molecule discussed earlier, particularly with regard to the axial or equatorial conformations in general and the hydroxyl group on carbon atom 1 in particular) it was decided temporarily at least not to introduce the extra-cyclic atoms as searchers. It may be noted here that the solution was actually achieved by using these atoms of orientation A later.

In order to gain further evidence on whether orientation A was correct or not the following method was adopted:

$$F_{hol} = 2\sum_i f_i \cos 2\pi(hx_i + lz_i)$$

where x_i, z_i are the atomic parameters.

Thus, if $x_i = X + \xi_i$, $z_i = Z + \zeta_i$,

$$F_{hol}(X,Z) = R_C \cos 2\pi(hX + lZ) - R_S \sin 2\pi(hX + lZ)$$

where $\begin{Bmatrix} R_C \\ R_S \end{Bmatrix} = 2\sum f_i \begin{Bmatrix} \cos \\ \sin \end{Bmatrix} 2\pi(h\xi_i + l\zeta_i)$.

Where R_C, R_S are calculated for a particular molecular orientation, F_{hol} may be calculated as a function of X, Z, the coordinates of the arbitrary origin used in computing R_C, R_S.

F is in any case limited by the inequality

$$|F| \leqslant \sqrt{R_c^2 + R_s^2}$$

R_c and R_s were calculated for all the F_{hol} for orientation A, and while, of course, the completely unknown position of the water molecule complicates matters, when its maximum contribution was added to the right hand side, the inequality was satisfied for 90% of the structure factors and the failures were, on the whole, not serious. This was counted as good supporting evidence for the orientation. Since for the reflexions where the inequality failed and also in cases where it was only just satisfied, $R_c \cos 2\pi(hX + lZ) - R_s \sin 2\pi(hX + lZ)$ must have very nearly its maximum value, each such reflexion imposes a restriction on the possible value of $hX + lZ$.

The maximum value is attained when

$$2\pi(hX+lZ) = \arctan\left(\frac{-R_s}{R_c}\right)$$

If these are plotted they should intersect at a point giving the coordinates of the molecule. When this was done a not unreasonable intersection was found, which was, in fact quite close to the true value; unfortunately at the time these methods were combined with similar techniques on reflexions which were very much smaller than $(R_c^2+R_s^2)^{\frac{1}{2}}$, where the choice of X,Z must be made in order to obtain a small enough value. Such reflexions are usually of low order and extinction errors in these made nonsense of the attempt and reduced confidence in the whole procedure. (It may seem paradoxical that extinction should affect those reflexions where F(X,Z) had to be made small while not those where it had to attain its maximum value, but this depends solely on the values of R_c and R_s : several of the |F| involved in the 'maximisation' procedure were quite small and those of the 'reducing' procedure among the largest).

Although it was not realised until later (after a conversation with Professor Lipson), R_c and R_s are the real and imaginary components of the Fourier transform of the molecule.

Dr T.R.R. MacDonald drew attention to the fact that the quantities R_c and R_s are those which are required for a method developed by Dr F. Stern and used by him to place the sucrose molecule in the cell (Beevers et al. 1952). It consists in principle in comparing Patterson projections for all values of X,Z with that derived from the observed data.

The Patterson function is given by

$$P(u,w) = \sum_h \sum_l F^2_{hol} \cos(hu+lw)2\pi$$

Now let $P'(u,w,X,Z) = \sum_h \sum_l F'^2_{hol}(X,Z) \cos(hu+lw)2\pi$

Then $\Phi = \iint PP' \, du \, dw$

$$= \tfrac{1}{2}\sum\sum F^2(R_c^2+R_s^2) + \tfrac{1}{2}\sum\sum F^2(R_c^2 - R_s^2)\times \cos 2\pi(hX+lZ) - \sum\sum R_c R_s \sin 2\pi(hX+lZ)$$

Φ is a function of X,Z which is positive and large when peaks in P' coincide with peaks in P and it will have its maximum value when all peaks in P' coincide with the corresponding peaks in P, i.e., when X,Z take their true valus.

In the space group $P2_1$ there are three types of non-zero vector:

$x_1 - x_2$, $y_1 - y_2$, $z_1 - z_2$ "Difference vectors"

$x_1 + x_2$, $\frac{1}{2} + y_1 - y_2$, $z_1 + z_2$ "Sum vectors"

$2x_1$, $\frac{1}{2}$, $2z$, "Harker vectors".

The Harker vectors are special cases of sum vectors.

The difference vectors are independent of the position of the molecule in the cell; expressed in terms of X, Z, ξ_i, η_i the coordinates of the sum peaks become

$$2X+\xi_1+\xi_2 \ , \ \tfrac{1}{2}+y_1-y_2 \ , \ 2Z+\eta_1+\eta_2$$

Let $\underline{R}$ be the vector from O to X, Z and $\underline{di}$ be the vector from (X, Z) to the i^{th} atm. The molecule is defined by the $\underline{di}$ which, if the chosen orientation is correct, also occur in the crystal.

The sum peaks are given by $2\underline{R}+\underline{di}+\underline{dj}$ in P' and by $2\underline{ro}+\underline{di}+\underline{dj}$ in P, if $\underline{ro}$ is the true value of the vector to the orbitrary origin. The maximum value of Φ should occur when $\underline{R} = \underline{ro}$ for then all sum peaks coincide (the difference peaks always coincide if the orientation is correct). However, Φ will attain a large value whenever

$$2\underline{R}+\underline{di}'+\underline{dj}' = 2\underline{ro}+\underline{di}+\underline{dj}$$

i.e. when $\underline{R} = \underline{ro} + \tfrac{1}{2}\{(\underline{di}+\underline{dj}) - (\underline{di}'+\underline{dj}')\}$

$= ro + \tfrac{1}{2}$ (a - a') say.

Since by hypothesis of the orientation is correct, $\underline{di}$ occurs both in the crystal and the assumed configuration, a and a' must occur in both sets (i, j) and (i', j') so that Φ will also be large when

$$\underline{R} = \underline{ro} - \tfrac{1}{2}(a' - a)$$

i.e. whenever $\underline{R} = \underline{ro} \pm \tfrac{1}{2}(a - a')$

so that a centre of symmetry at $\underline{R} = \underline{ro}$ should occur. This would seem to provide an additional means of identifying $\underline{ro}$ and it was actually this fact that was conclusive in the work on sucrose, where several large peaks of much the same height occurred.

However there may occur coincidences between the sum peaks of P' and the difference peaks in P (given by $(\underline{di} - \underline{dj})$. This may produce large values of Ψ when

$$2\underline{R} + \underline{di}' + \underline{dj}' = \underline{di} - \underline{dj}$$

i.e. when $\underline{R} = \tfrac{1}{2}\{(\underline{di} - \underline{dj}) - (\underline{di}' + \underline{dj}')\}$

$$= \tfrac{1}{2}(b - a')$$

These positions in general do not have a centre of symmetry although this is not strictly true for cases where there are a large number of multiple vectors, as in a sugar. They certainly have no symmetry about $\underline{ro}$ however and so these undesirable coincidences will detract from the useful symmetry of the synthesis.

In computing the Ψ synthesis the origin peak in P was subtracted, for the large size of this might well cause many spurious and confusing regions of high value. The resulting map showed a remarkably good approximation to a

centre of symmetry, but unfortunately this did not coincide with the maximum which was about 0.8 A away. Both positions were investigated by calculation of the one dimensional structure factors, the water molecule being placed to give as good agreement as possible, but the agreement was not particularly good and the packing, from the hydrogen bond point of view was distinctly poor. The structures were therefore provisionally abandoned. It is interesting to note that that based on the maximum of the synthesis is about 0.25 A from the correct structure; its rejection was due primarily to the extinction errors in F_{hoo} which caused the water molecule to be placed in an entirely wrong position where the packing was poor, if not impossible. The reasons why a centre of symmetry should have occurred in the synthesis at quite the wrong place are not understood; the undesirable interactions of sum and difference peaks, coupled with complete ignorance of the effect of the water molecule in P' make a complete failure of the method not beyond conception, but the spurious centre of symmetry remains very odd.

An attempt was made to place restrictions on the position of the molecule by packing

considerations alone. By drawing spheres whose radii are the Van der Waals diameters round the carbon atoms and the standard hydrogen bond lengths as diameters round the oxygen atoms, the permissible parts of the unit cell in which the atoms of the other molecule must lie can be seen. Since the separation of symmetry-related atoms is 2x, $\frac{1}{2}$, 2z, the position of the screw with respect to one atom must be such that its mate lies in a permissible site on a section $b/2$ above it. By using these considerations for each atom in turn it is possible to limit the area in which (X,Z) must lie in order that the molecules can pack in an acceptable fashion. The result of this was quite a small area in appearance, but large enough to make purely trial and error methods a formidable undertaking. Some doubt was also felt as to whether the area necessarily included the true value of (X,Z) since when many atoms are involved, slight changes in the model adopted for the method may cause relatively great changes in the permissible area.

The results of the Fourier transform calculations were sufficiently encouraging for confidence to be felt that the orientation of the molecule was correct and in this belief

the extra-cyclic atoms of the ring were introduced as further searches in the Patterson superposition technique. By this means a large number of spurious atomic sites was eliminated. Since two of the oxygen atoms (O_3, O_5) occur at the same y level the device of searching sections $b/2$ apart for the same vector between peaks, as mentioned above, was systematically explored. This gave rise to only two possibilities, one of which placed the screw axis in an obviously wrong position; the other gave promising sites for most of the other atoms of the second molecule and more careful examination of the Patterson function revealed that the dubiety of the other sites wasmainly due to the low level of the carbon Harker peaks and that by discounting these, excellent agreement with the superposition maps could be obtained. Search of sections of the maps $b/2$ apart for two further sites related by this screw axis gave only one position for the water molecule. An examination of the packing of the molecules in this structure showed no unreasonably short distance between non-bounded atoms and indicated that six hydrogen bonds held the structure in a three-dimensional network; this is the number of hydrogen atoms which are available for bonding. In order to check the correctness of

this structure (which was the nineteenth trial structure proposed), all the vectors between atoms were calculated and plotted on the three-dimensional Patterson function. The agreement was excellent and all the features of the function were adequately explained. The first value of the reliability factor for the hol reflexions was R = 33%.

The candidate has been able to draw several morals from the long and somewhat tangled pattern of events which led to the solution of the structure. The first is that too much reliance should not be placed in the matching of a few structure factors by trial and error, particularly in low orders where extinction effects may be serious. There can be no doubt that extinction, particularly of F_{100} (whose observed value remains anomalously low even after extinction corrections have been applied) delayed the solution of the structure for more than a year. F_{100} was a particularly unfortunate example to be affected by extinction, for, quite apart from the psychological damage inflicted by violent disagreement in what, for most trial structures, was the first structure factor to be calculated, it is one which is ideally suited for trial and error location of the molecules in the cell - the value of the

molecular Fourier transform at this point is so very large that the position of the molecules is quite critical in achieving reasonable agreement; and quite generally, low wave numbers, being comparatively insensitive to small changes in atomic positions in the trial molecule, and also in giving fewer possible positions in the cell, are to be preferred for this purpose. A further lesson which may be drawn from this is the value of the Patterson function, which contains information from all the diffraction data; even large errors in a few of these make very little noticeable difference in the appearance of this, which is thus a very trustworthy tool in the solution of structures. Another great advantage of it is that parts of a crystal structure can be handled at a time which is not generally the case in methods that depend on the calculation of structure factors, unless heavy atoms are involved.

Refinement of the Structure

The b axis projection was refined first, since only in this projection is there an effective centre of symmetry. Initially electron-density projections and least-squares methods were used but latterly ($F_o - F_c$) syntheses, which are formally equivalent to least squares techniques with the off-diagonal terms of the matrix ignored and with somewhat arbitrary weighting factors, were solely employed. This method has the advantage over the ordinary least squares procedure of showing up significant discrepancies in the 'observed' and calculated electron densities which may be due to hydrogen atoms not included in the structure factor calculations and effects of thermal vibrations of the atoms; it also lends itself to the efficient exploitation of the computing facilities available in this laboratory.

The refinement was first carried out without removing the hydrogen atoms with the same temperature factor ($B = 2.4\ A^2$) applied to both oxygen and carbon atoms. In a structure such as this the contribution of the 14 independent hydrogen atoms, whenever the majority of them scatter in phase, can be of significant proportions but as only five (those of the CH

groups) can be expected to occur in predictable positions, placing the others was not easy. In the b-axis projection, overlapping is not serious as far as the heavy atoms are concerned (with the exception of O_4 and O_6), but many of the hydrogen atoms were expected to fall near the heavy atoms; this made interpretation of the difference maps very difficult, since high electron density near an atom might mean a shift, a thermal vibration of the atom, or the presence of a hydrogen atom, or indeed, a combination of all three. When the five hydrogen atoms of the CH groups were introduced into the structure factor calculations, by placing them about 1A from the carbon in a direction such as to make the tetrahedral angle with the other bonds made by the carbon atom, some improvement in agreement was observed and the interpretation of the difference maps made a little easier. At this stage the evidence for the positions of the hydrogen atoms in the hydrogen bonds was inconclusive; of these atoms, one, that in the bond between O_3 and O_5, must belong to O_3, since O_5 is the ring oxygen. The others are arranged in two sets; one consists of the two hydrogens in $O_4 - O_6$, $O_4 - O_6'$ and the other of the three

in $O_1 - O_6$, $O_6 - O_2'$, $O_2' - O_1$. If the position of one in either set is fixed, the positions of the others in the set must also be fixed, but the two sets are independent of each other. A statistical distribution throughout the crystal is not, of course, impossible or even unlikely. The evidence for these positions was, as mentioned, inconclusive if not contradictory. The hydrogens of the methyl group may be locked in position, or oscillating or rotating and while considerable electron density was shown on the maps it was impossible to reach any conclusion about their positions either; to add to the complication, it was suspected that the carbon atom of the methyl group was indulging in considerable anisotropic vibration. The discrepancy factor R for the F_{hol} was at this stage 15.4%.

The a axis projection was then refined in a similar fashion. The R factor was reduced from 26% to 11.6% by five successive cycles and the position achieved was then very similar to that of the b-axis refinement. Although the agreement was better, this was to be expected in a non-centrosymmetric projection and the overlapping in this case was much more serious. Some bond lengths were calculated from the atomic

parameters obtained; the results were very odd, C - C lengths for instance, which would not be expected to differ significantly from 1.54 A, had a large scatter and ranged from 1.40 A to 1.65 A. It was realised that the structure was far from accurate and that the only way to achieve much improvement, short of three dimensional computations for which this laboratory is not well equipped, would be to employ as accurate intensity data as could be obtained. The reasons for the poor accuracy of the structure probably lie in the lack of resolution of the projections, coupled with the non-centrosymmetric character of all but one of them. These circumstances are liable to occur, of course, whenever a substance is studied for the intrinsic interest of its structure rather than because of possible crystallographic elegance.

The collection of intensity data using the Geiger counter has been described above; only two zones were investigated, the rectifier burning out on completion of the second and a replacement arriving too late for the collection of the I_{okl} for this work. The F_{obs} so obtained were scaled directly with the F_{calc} of the previous structure; it is interesting to note

that the average temperature factor B for the Geiger counter structure factors was considerably lower than that obtained by Wilson's (1942) method on those visually estimated (1.1 A^2 in place of 2.4 A^2). This is thought to be a systematic effect which may apply generally to visually estimated intensities.

Refinement of the F_{hol} was resumed with these data. R showed a slight initial rise, showing that the previous F_{hol} had been over refined and then fell monotonically. The difficulties mentioned above due to overlapping were still present and a measure of trial and error had to be introduced in some cases by altering a parameter and examining the subsequent difference map to see if the electron density distribution at the atom in question justified the change. Individual isotropic temperature factors were introduced and it was possible to choose, although somewhat tentatively, sites for the hydrogens engaged in bonding and, later still, for the methyl hydrogens. The R factor was at length reduced to 7.8% and the map showed no great evidence for further parameter changes; it was obvious that some anisotropy was present in the thermal vibrations, however.

The c-axis projection was then refined and here again great difficulties due to overlap were encountered. The most startling feature of the first difference synthesis was a large and unequivocal shift in the x-parameter of O_3, which is well resolved on the b axis projection (apart from hydrogens) and while it is vibrating anisotropically, it was thought that its coordinates had been determined with reasonable accuracy. While on the c axis projection O_3 is poorly resolved from C_3 and O_5 (with the additional complication of the hydrogen in the O_3 - O_5 bond) there was no doubt that in order to achieve the best agreement with the F_{obs}, a shift had to be made. This point will be raised again later.

Refinement proceeded much as before; some evidence again was gained for the positions of the hydrogen atoms which led to them being placed in the same positions as had been found in the b axis refinement, but on the whole the evidence was not as strong as could be wished for. It is particularly important in the case of a non-centrosymmetric structure that hydrogen atoms should be included as early as possible in the refinement (due regard being taken to the correctness of their placing) for otherwise not

only are the moduli of the calculated structure factors wrong, but their phases also. The difference that the hydrogen atoms can make to the phases, and thereby the difference maps, is discussed later. The refinement was continued until R = 6.1% and the x parameters from the two refinements were compared. With the exception of O_3 the x parameters of the atoms from the two independent refinements agreed closely; the maximum difference was .004 of the cell edge which applied only to one atom which was unresolved in the first projection. Means were taken, the centrosymmetric projection being preferred in cases where the atoms were reasonably resolved on it. In the case of O_3 there was a difference of .007 and the centrosymmetric value was therefore increased by .002; this weighting seems reasonable when the diminished accuracy of the non-centrosymmetrical projection in general is coupled with the poor resolution of this atom on it. The mean values thus obtained were used as the atomic coordinates of the final structure.

Before the final structure is discussed, it is thought advisable to give some details of the actual procedure used in the final stages. The F_{obs} were scaled agains the F_{calc} after each

cycle by summing them in ranges of $\sin^2\theta$, plotting the ratios and taking the best straight line; the value of the ordinate at the value of $\sin^2\theta$ for a particular reflexion was then used as a factor to scale F_{obs}. This means, of course, that the values of the temperature factor B listed later have no absolute significance. It was noted that the largest values of F_{calc} were always considerably in excess of F_{obs}; this was attributed to secondary extinction and corrections were accordingly made by the formula $I' = \frac{I}{1 - gI}$. The factor g was chosen so as to give the best agreement for the largest values of I but as this correction always increases ΣF_{obs}, the structure factors had immediately to be scaled, which reduced F_{obs} once more, so that the choice of g had to be made by successive approximations. This correction made agreement of the large F satisfactory with the exception of F_{100}, which, not having one of the very largest intensities, received only a moderate boost from the procedure. It will be recalled that the low value of this structure factor played an important part in the delay which attended the solution of the structure. The observed value is remarkably constant, if the effect is due to primary extinction; photographs

of four crystals of varying size, grown in three batches under dissimilar conditions and subjected to cutting and rolling gave the same value to within 5%, and the Geiger counter measurements, where indeed this was the standard used so that there is no possibilityof a gross error in measurement (such as miscounting the foils), gave a result within the range. The effect must presumably be ascribed to primary extinction; in any case, this structure factor has been omitted from the difference synthesis and error calculations, and its calculated value used in the electron density projections.

The atoms were placed to the nearest .001 of the cell edges and two-figure tables used to derive their trigonometric functions. The various maps were computed on the Stern analogue machine at intervals of $^1/_{30}$ of the cell, the coefficients of the difference syntheses were multiplied by a factor usually 20, to reduce computing errors. The scattering factors of Berghuis et al. (1955) were used.

The percentage discrepancy factors for the two zones of structure factors where Geiger counter data were available, using the parameters given in the table below, were

$$R_{hol} = 7.6\%$$

$$R_{hko} = 6.3\%$$

F_{100} is excluded from both these summations; it is discussed later. Values of F_c where F_o is not observed are included at their full value.

The mean coordinates and temperature factors derived are given in the following table.

	x/a	y/b	Z/C	$B_{hol}(A^2)$	$B_{hko}(A^2)$
O_1	.188	.726	.110	2.4	1.5
O_2	.880	.040	.225	0.8	1.0
O_3	.060	.007	.605	1.5	2.0
O_4	.358	.182	.504	0.8	1.0
O_5	.187	.002	.016	1.5	2.0
O_6	.328	.533	.460	2.0	1.5
C_1	.087	.867	.080	2.0	0.8
C_2	.008	.906	.272	0.5	0.3
C_3	.143	.976	.431	0.5	2.4
C_4	.232	.130	.353	1.0	1.5
C_5	.315	.079	.159	1.5	1.5
C_6	.402	.224	.059	2.4	2.4

The large variation in the temperature factors of C_1 and C_3 in the two projections was unexpected. When anisotropy in the thermal vibrations is present there is, of course, no reason why they should be the same, since the

appearance of such an atom as a scatterer of X-rays will obviously depend on direction, but such a large variation, particularly in atoms belonging to the ring is a little disturbing. The atomic coordinates of the hydrogen atoms are given in the following table

	x/a	y/b	z/c	all in $\frac{1}{60}$'s of the cell edge
H_1	0	48	59	
H_2	58	47	20	
H_3	14	53	29	
H_4	8	14	18	
H_5	24	1	12	
H_6	26	20	11	The methyl group
H_7	20	18	59	
H_8	33	15	2	
H_9	6	1	44	
H_{10}	21	16	29	
H_{11}	26	35	28	
H_{12}	10	40	0	
H_{13}	16	37	20	
H_{14}	11	32	40	

It must be emphasised that these values were assumed, from a certain amount of experimental evidence, for the purpose of assessing the hydrogen contributions to the structure factors; they should not be taken as accurate experimentally derived parameters for the purpose of calculating C - H and O - H bond lengths. The evidence for the actual hydrogen positions will be discussed later. The scattering factor for the hydrogen atoms was unmodified by any temperature factor: it was found that the change caused by any reasonable value of a temperature factor was hardly appreciable.

The calculation of the structure factors of the final structure gave data for the assessment of errors in the electron density and the atomic coordinates by the method of Cruickshank (1949).

The values of $\sigma(\rho)$ for the two projections, given by

$$\sigma(\rho) = \frac{1}{A}\left\{\Sigma(F_o - F_c)^2\right\}^{1/2}$$

the summation being over the complete zone of structure factors, gave

$$\sigma(\rho_{h0l}) = 0.245\ e A^{-2}$$

$$\sigma(\rho_{hk0}) = 0.36\ e A^{-2}$$

the value of $\sigma(\rho_{hko})$ here having been multiplied by two because the projection is not centrosymmetrical (Cruickshank 1950). In these calculations when F_{obs} is zero, what is estimated as the highest value which could remain unobserved has been used to give $(F_o - F_c)$; this applies also to the difference of syntheses. The effect of this is naturally most marked in the centrosymmetrical projection, where, if unobserved terms are omitted completely $\sigma(\rho)$ falls to 0.225 A^{-2}. The corresponding standard deviations in the atomic parameters given by

$$\sigma(x) = \frac{2\pi \left\{\sum h^2 (F_o - F_c)^2\right\}^{1/2}}{a A C}$$

where C is the central curvature of the atom in question, are given in the following table.

		b axis	a axis
Oxygen	$\sigma(x)$	.0075	-
	$\sigma(y)$	-	.018
	$\sigma(z)$	.0064	-
Carbon	$\sigma(x)$	.011	-
	$\sigma(y)$	-	.020
	$\sigma(z)$	.010	-

These values have been multiplied by the factor $\left(\frac{u}{u-n}\right)^{\frac{1}{2}}$ where u is the number of independent F_{obs} values and n the number of parameters to be refined (Cruickshank 1949)

which is here in each case 36, two coordinates and a temperature factor for each of the 12 atoms. This factor has the values 1.27 and 1.60 for the hol and hko refinements respectively. In addition, the value of $\sigma(y)$ given above includes the factor two because of the lack of centre of symmetry. This factor in this structure should ideally be less than two because of the occurrence of the real F_{hoo} in the non-centrosymmetric (hko) projection but the difference will be small in comparison with the uncertainty in measurement of the curvatures of the atoms and is ignored; the effect of this is in any case on the safe side in the estimation of the standard deviation of atomic parameters. The curvatures of the atoms were measured from the electron density projections shown in Figures VI and X. The derivations were made from the approximation

$$\rho(r) \propto \exp(-pr^2)$$

whence

$$\left(\frac{\partial^2\rho}{\partial r^2}\right)_{r=0} = -2p\rho(0)$$

(Lipson and Cochran 1953, p. 301)

the origin of r being at the centre of the atom in question.

The value of p was found from the slope of the line given by plotting $\log\rho$ against r^2 and $\rho(0)$ was deduced from the electron density

projections. Some difficulty was encountered in doing this, particularly in the c-axis projection where none of the carbon atoms are properly resolved but the results should not be in great error. The curvaturesshould ideally be evaluated for each atom separately since in general they will differ, the values depending on the thermal vibrations; this was impossible here, because of the bad resolution and the following values, which are averages over several atoms in each projection were used

	Oxygen	Carbon
c axis projection	$-110\ eA^{-4}$	$-95\ eA^{-4}$
b axis projection	-140	-100

The various maps obtained from the final structure may now be discussed. The electron density projections can be disposed of first, since apart from their use in estimating the curvatures of the atoms, their only real value is to give a general picture of the structure for the benefit of non-crystallographers. The b-axis projection (Fig. VI) gives a reasonably constant height of about $14eA^{-2}$ for the oxygen atoms; the carbon atoms (as far as can be measured) fall into two groups, C_3 and C_4 having heights of about 10.4 eA^{-2}, the

others about 9 eA^{-2}. This difference, which would appear to be significant is not fully explained by the differing thermal vibrations of the atoms; the correlation between peak height and temperature factor is not particularly striking. The effects of termination of the series are probably involved in this phenomenon; the positions of the maxima in the electron density distribution, which do not give accurately the coordinates used in the calculations, show that these effects are serious, which, owing to the low limit of which was attained with the Geiger counter, is hardly surprising.

The electron density projection down the c axis shows similar features with the oxygen peak height reasonably constant at about 12.5 eA^{-2} and the carbon peak heights varying from about 8 to about 9.5 eA^{-2}; C_4 has the lowest value here, in contradistinction to the (hol) projection. These electron density maps were calculated with less care and accurary than would be necessary if they were to be used as the basis of a refinement procedure and it is felt that this, coupled with termination of the series effects can explain away these discrepancies.

The difference synthesis, with all atoms removed, of the projection down the b-axis (Fig. IV) reveals as its main features the anisotropic vibrations of several atoms, notably O_1, O_6 and C_5. O_3 is thought to be engaged in more active anisotropic vibration than is indicated since the coordinate change discussed above, which is shown here as a small but definite indication to move back again, tends to mask this. Anisotropic corrections to the structure factor calculations were not made because of the labour involved; while in principle if the temperature factors used are not correct, the parameters obtained from $(F_o - F_c)$ synthesis are not entirely free from termination of the series errors, it is felt that in this case they will not be serious since the electron density involved is small; in any case when the overlapping nature of the electron-density projections is recalled, it would obviously be impossible to make this correction accurately by comparing F_o and F_c syntheses.

A difference synthesis with the contributions from the hydrogen atoms excluded from the F_c is also shown (Fig. V). Although the resolution is far from good it can be seen that the

general features of the map support the locations adopted. The hydrogens of the methyl group appear with moderatedefinition, although since they are superposed on the typical distribution due to anisotropic vibration of the carbon atom, their appearance must be treated with caution. The atom in the bond O_1 - O_6 has the poorest definition, but as it is superposed on one of the troughs due to the anisotropic vibration of O_6, this is not unreasonable. There is no guarantee, of course, that the maximum of the hydrogen electron distribution will lie on the line joining the oxygen atoms which it bonds, and indeed in salicylic acid (Cochran 1953) it is displaced from this line. Such changes in the assumed positions of the hydrogen atoms could improve the agreement of the postulated positions and the maps, but the evidence is not sufficiently reliable to warrant this move. The peak heights of the hydrogen atoms are very variable, but if account is taken of the background on which they lie, a height of about $1eA^{-2}$ is a reasonable average.

Attention may now be given to the c-axis projections. Figure VII shows the $(F_o - F_c)$ synthesis with all the atoms removed. The most striking feature is the gradient on which

O_3 lies. This is due entirely to the compromise which had to be adopted in assigning the x-coordinate of this atom. Different maps, in which the value of x, deduced from the independent refinement of this projection was used for the structure factor calculations, were quite flat at O_3 and it was only after the comparison of the two projections gave the final parameters, that this feature appeared. In particular, what might be held on this map to indicate a change in the y-coordinate of C_3 was previously absent. It may be noted that the actual change indicated in x_{O3} corresponds to 0.025 - 0.030 A, which when the standard deviation of the position on an oxygen atom in this projection is recalled (σ = .018), is not significant on the levels of significance suggested by Cruickshank (1949). Since the standard deviation of the electron density in this projection is 0.36 eA^{-2} most of the detail shown on this map may well be spurious, although it is possible to interpret parts of the distribution as being caused by anisotropic thermal vibrations.

In refining non-centrosymmetrical projections by difference maps, where the coefficients of the syntheses instead of

$(F_o - F_c)$ are $(A_o - A_c)$ and $(B_o - B_c)$, the A_{obs} and B_{obs} are usually found by the formula

$$A_{obs} = Ac. \frac{Fo}{Fc} \qquad B_{obs} = Bc. \frac{Fo}{Fc}$$

This ensures that $A_o^2 + B_o^2 = F_o^2$ and that the phases of the structure factors are preserved. For the refinement of atomic parameters, provided the n-shift rule is used, this works well, but in searching for hydrogen atoms, a difficulty of principle is raised: if the positions of the hydrogen are unknown, so that their contributions to the A and B cannot be calculated, the phases of the structure factors will be wrong even if the heavy atom parameters are correct and the resulting synthesis cannot possibly give the correct electron distribution. Cochran (private communication 1954) gives the opinion that the atoms should appear in their correct positions, but having only half the peak height, but it was found that in practice, by examining a hypothetical one-dimensional structure with the order of overlapping that occurs in rhamnose, some changes in the atomic position can occur as well. This knowledge must make the location of hydrogen atoms whose positions are quite unknown a rather uncertain procedure. Moreover

it is felt that if positions of hydrogen atoms are assumed and difference syntheses computed with the phases of the complete (but as far as the hydrogen positions are concerned somewhat hypothetical) structure, these phases will tend to produce electron density in the assumed positions whether they are in fact occupied by hydrogen atoms or not. This is an extension of the well known principle that, given reasonable agreement between the calculated and observed structure factors, Fourier maps will tend to yield what is put into the calculation of their coefficients; this is particularly true in non-centrosymmetric cases.

The positions of the hydrogen atoms in the rhamnose structure are not completely unpredictable. As discussed above, the uncertainty lies principally in the orientation of the methyl tetrahedron and in the choice between the alternative positions of the two sets into which five of the six hydrogen bonding atoms can be divided. The evidence of the centrosymmetrical projection seems to favour quite strongly the sites which were adopted, but since the final stages of refinement of the heavy atoms were carried out with these atoms removed and since many of the

possible hydrogen sites are confusingly superposed on effects interpreted as evidence for thermal vibration, it is possible that this evidence may not be unequivocal; in any case independent corroboration would be very desirable.

Figure VIII shows a map of the difference synthesis projected down the c-axis, using the phases given by the heavy atoms only i.e. with incorrect phases. The assumed positions are, on the whole, supported by the appearance of reasonably high electron density, although the maxima do not coincide very convincingly with these positions. The alternative positions for the two sets are of particular interest. It is a little distressing that of the set $H(O_4 - O_6)$, $H(O_4 - O_6')$ quite high values of electron density do occur near both sets of positions and while the positions of these satisfy the assumed set considerably better, the conclusions reached above as to the reliability of the positions make any definite assignment a little dubious. The evidence of the (hol) data on this set is favourable to the assumed position for one and neutral for the other since the resolution of O_4, O_6 is very bad. Of the other set $H(O_1 \leftrightarrow O_6)$ would fit a little better in the alternative position,

although it is quite passable in its assumed locations; $H(O_1 - O_2')$ and $H(O_2 - O_6')$ prefer the assumed positions. The hydrogens of the methyl group fit the map well enough, although other orientations would obviously be possible if a little deterioration were permissible.

Figure IX shows the difference synthesis in which the phases used were those of all the atoms, the hydrogens being placed in their assumed positions for the calculation of structure factors. As was expected, the assumed positions are now almost violently preferred, the only atom apparently rather poorly placed being one of the methyl group.

Great caution must obviously be used in making any decision about the positions of the hydrogen atoms which are not predictable on chemical grounds. The evidence of this work would indicate that the most probable positions are those which have been listed above, but that in the case of methyl group other orientations might well be possible, and that in the case of the atoms taking part in hydrogen bonding, a statistical distribution is not impossible.

As was explained above, it was not possibe to obtain Geiger counter data for the a axis

projection. This projection had been used in the first refinement using visually estimated intensities, however, and it was obviously of interest to see how these data agreed with the final structure. The zone of structure factors was therefore calculated from the coordinates of the final structure, but omitting the hydrogen atom contributions and using only the average temperature factor $B = 2.4A^2$, in order to conform with the conditions under which the previous calculations had been carried out. In spite of considerable differences in the atomic coordinates (which were sufficient to reduce the abnormal values of the original bond lengths to those given later) the overall agreement remained almost constant with $R = 11.8\%$. This adds weight to the opinion formed that the refinement of non-centrosymmetric projections requires critical examination and that it would require more than a low discrepancy factor to inspire confidence in the accuracy of such a structure. What is most disquieting about this is that the former refinement seemed to be at, or at least near, an end as far as changes in atomic coordinates were concerned. It was not fully realised

at the time that the inclusion of hydrogen atoms and the assignment of individual temperature factors could result in considerable changes of these coordinates and a natural diffidence was felt about introducing hydrogen atoms without any really convincing evidence for their positions.

This experience emphasises the dilemma outlined above that in a poorly resolved non-centrosymmetrical projection the contributions of the hydrogens are necessary for the refinement to proceed smoothly, but that until a fairly advanced stage of refinement is reached, their positions may remain unknown. This argument reinforces the need for the highest possible accuracy in the diffraction data, since it is by using such measurements that the hydrogen positions are likely to be found at the earliest stage.

The values of F_o and F_c will be found in an appendix.

The bond lengths between atoms in the molecule are given in the following table

$C_1 - O_1$	1.376 A
$C_2 - O_2$	1.478
$C_3 - O_3$	1.410

$C_4 - O_4$	1.408
$C_5 - O_5$	1.456
$C_1 - O_5$	1.420

$C_1 - C_2$	1.512
$C_2 - C_3$	1.534
$C_3 - C_4$	1.523
$C_4 - C_5$	1.564
$C_5 - C_6$	1.528

The hydrogen bond lengths between adjacent molecules are:

$O_1 - O_2'$	2.69 A
$O_1 - O_6$	2.91
$O_2 - O_6'$	2.79
$O_3 - O_5$	2.83
$O_4 - O_6$	2.80
$O_4 - O_6'$	2.74

A formula for the standard deviation of a bond length $\sigma(l)$ in cases where

$$\sigma(x) \neq \sigma(y) \neq \sigma(z)$$

has been given by Ahmed and Cruickshank (1953). $\sigma(l)$ has been evaluated by this means only for those bonds in which a significant deviation from the accepted values (taken as 1.542 for C - C and 1.430 for C - O)

looks possible. Throughout this work, standard deviations have been computed in terms which are strictly applicable to orthogonal coordinate systems only: the assumption that β (= 95°37') is sufficiently close to 90° is not likely to alter the possible significance of the values of the molecular parameters found.

For the C - C bonds, the least possible value of $\sigma(l)$ is 0.014 (in the special case when the bond lies parallel to the c axis) so that the least possible deviation from the standard value which can be counted significant is $1.645\sigma \approx 0.023$ A, so that the only deviation which can possibly be significant is in $C_1 - C_2$. This bond length is to a large extent independent of $\sigma(y)$ and its standard deviation is 0.015 so that $\frac{dl}{\sigma} = 2.0$. The difference, therefore, on Cruickshank's levels, is possibly significant.

Similar considerations apply to the C - O ~~atoms~~ bonds and the deviations from the value of 1.430, which is a mean deduced from the tables of Allen and Sutton (1950) are certainly not significant except for $C_1 - O_1$ and $C_2 - O_2$ the axial bonds of the molecule. These bonds, which in the ideal model of the molecule are parallel will, with the approximations adopted here of assuming an average value for the

central curvatures of the atoms, will have the same standard deviation. This is calculated as 0.027 so that $\frac{dl}{\sigma}$ is 2.00 and 1.78 for $C_1 - O_1$ and $C_2 - O_2$ respectively. Both of these again fall in the range described as possibly significant.

The bond angles in the molecule are shown in the following table.

$O_5C_1C_2$	113.0°
$C_1C_2C_3$	110.1
$C_2C_3C_4$	111.5
$C_3C_4C_5$	108.8
$C_4C_5O_5$	109.1
$C_5O_5C_1$	120.1
$O_5C_1O_1$	108.7
$C_2C_1O_1$	109.0
$C_1C_2O_2$	103.8
$C_3C_2O_2$	107.7
$C_2C_3O_3$	106.5
$C_4C_3O_3$	114.2
$C_3C_4O_4$	107.9
$C_5C_4O_4$	110.1
$C_4C_5C_6$	114.2
$O_5C_5C_6$	109.9

The minimum standard deviation $\sigma(\varphi)$ which would occur when all three atoms have the same value of y was estimated by Ahmed and Cruickshank's (1953) formula as about 0.7°; on the other hand if the y parameters are largely involved it can rise to about 1.7°. The standard deviations of each angle have not been calculated but it is completely safe to say that the deviations of the following from the tetrahedral angle of 109.5° are significant, even assuming that $\sigma \sim 1.7°$.

$C_5O_5C_1$	120.1 °
$C_1C_2O_2$	103.8
$C_4C_3O_3$	114.2
$C_4C_5C_6$	114.2

In addition it is again safe to say that the following are possibly significant

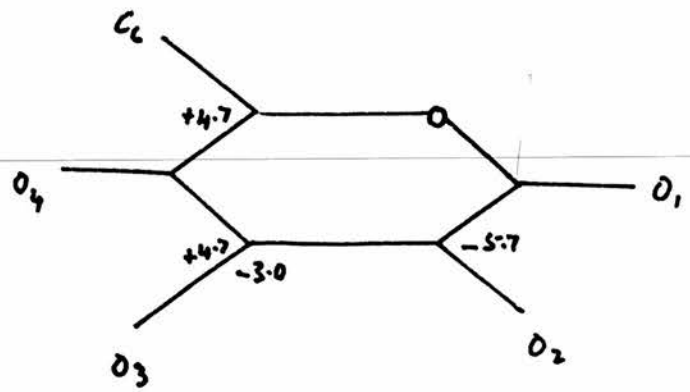
$O_5C_1C_2$	113.0 °
$C_2C_3O_3$	106.5

and that by consideration of the y parameters involved are almost certainly significant. The other variations are probably not significant.

A somewhat surprising feature is the angle of 120° at the oxygen atom of the ring. While the angle which bonds make at oxygen

atoms is not a well defined quantity in structural chemistry as is, for example, the tetrahedral configuration of a carbon atom making four single bonds, it has been generally assumed that in the pyranose ring this angle would have the tetrahedral value; in glucose (MacDonald 1950) the value of 111° was found which was not taken to differ significantly from 109.5° (no standard deviations are quoted).

The variations in the angles made by the extra-cyclic atoms seem to indicate a tendency for these groups to move away from each other and into the space surrounding the oxygen atom of the ring. This may be shown schematically as under; it must be remembered that $C_1 - O_1$, $C_2 - O_2$ are really pointing in opposite directions, at right angles to the 'plane' of the ring (as previously defined) i.e. at right angles to the plane of the paper.



The deviations from the tetrahedral angle are shown in degrees. This would seem a reasonable tendency which could be explained by simple repulsion of the substituents: O_3, O_4 and C_6 all occupy equatorial positions and while O_2 is axial, it is cis to O_3. No overcrowding will result from what appears in this purely schematic diagram to be an approach of O_2 towards O_1 since they are axial, and trans to each other.

A comparison of the structure of rhamnose with that of the glucose molecule may be made. This is hampered to some extent by the lack of information on the accuracy of the glucose structure - the limits of error suggested are almost certainly too low. The two features of striking interest about the glucose molecule were the two short C - O bonds; one of them, that occurring in CH_2OH group has no analogue in rhamnose, the other, $C_1 - O_1$ was found to 1.31 A in length. It seems almost certain that such a low value must be significant, although MacDonald was gravely hampered by poor resolution in his refinements, and it is gratifying to note that this bond in rhamnose has the possibly significantly low value of 1.37_6. It is tempting to think that this may be a feature of sugars in general, or

at least in those were the $(OH)_1$ group is axial, and while such a generalisation must await confirmatory evidence from work on other sugars, the coincidence is striking. The other bond length whose value is possibly significantly different from the standard value in rhamnose has no analogue in glucose where $(OH)_1$ is the only axial group. There is no experimental evidence of a systematic deformation of the bond angles of the extra-cyclic atoms in glucose.

The packing of the molecules and the water molecules in the rhamnose crystal is an excellent example of hydrogen bonding in three dimensions, in which all the available hydrogen atoms are used. A particularly interesting feature is the occurrence of a 'helical' chain round one of the screw axes involving O_4 and the water molecule ($-O_4-O_6'-O_4'-O_6-O_4-$); another notable point is the linking of molecules which repeat by translation along the axis by the bond $O_3 - O_5$. The angles which the hydrogen bonds make with the covalent bonds have not been calculated but appear from a model of the structure to be reasonably tetrahedral; in particular the disposition of the four hydrogen bonds round the water molecule is recognisably so. The length of

the bond $O_1 - O_6$ (2.91A) has an unusually high value; but one of length 2.93A, again linking a hydroxyl group and a water molecule has been reported (Pitt 1948).

In this thesis, as indeed in the work which led to its presentation, emphasis has been laid on the accumulation of reliable data rather than on the formation of theories; the reason for this is the necessarily speculative character of the latter in the absence of the former, and in the field of structural carbohydrate crystallography data of any type are singularly scanty.

After reading the discussion of this work, the examiner may feel that the candidate regards his work with a sense of shame; this is not so and no apology is offered for concentrating on the dubious features of the work rather than ignoring them and stressing its merits, which the candidate feels, should be obvious. The striking agreement of the structure factors and the low level of electron density on the final difference maps, which have been achieved under circumstances of some difficulty, speak for themselves; but it has happened so often in scientific research in general and crystallographic research in

particular, that deductions have been made from data whose accuracy does not warrant it, that the only scientific approach to the matter was to discuss openly any doubtful features and to try to ensure that facts which can be drawn from the study can be used with confidence.

ACKNOWLEDGMENTS

The candidate wishes to place on record his great debt to Dr C.A. Beevers, whose continual advice and encouragement has been invaluable. He also wishes to acknowledge with gratitude financial support from D.S.I.R. and the Carnegie Trust for the Universities of Scotland. Some of the apparatus needed was bought with funds supplied by the Agricultural Research Council and the Moray Fund of the University of Edinburgh.

REFERENCES

Ahmed and Cruickshank (1953). Acta Cryst. 6, 385.

Allen and Sutton (1950). Acta Cryst. 3, 46.

Andress and Reinhardt (1931). Zeit. f. Krist. 78, 477.

Beevers (1952). J. Sci. Inst. 29, 230.

Beevers and Cochran (1947). Proc. Roy. Soc. (A) 190, 257.

Beevers, MacDonald, Robertson and Stern (1952). Acta Cryst. 5, 689.

Berghuis, Ijbertha, Haanappel, Potters, Loopstra, MacGillavry and Veenendaal (1955). Acta Cryst. 8, 478.

Boeseken (1949) in 'Advances in Carbohydrate Chemistry' 4.

Charalambous (1953) Thesis. University of Edinburgh.

Cochran (1948). Acta Cryst. 1, 138.

Cochran (1950). Acta Cryst. 3, 268.

Cochran (1953). Acta Cryst. 6, 260.

Cox (1931) J.C.S. 2313.

Cox (1932) J.C.S. 2536.

Cox et al. (1932) J.C.S. 1844.

Cox et al. (1935) J.C.S. 978, 1495.

Cox et al. (1936) J.C.S. 769.

Cox and Jeffrey (1939). Nature 143, 894.

Cruickshank (1949). Acta Cryst. 2, 65; 2, 154.

Cruickshank (1950). Acta Cryst. 3, 72.

Eeland and Pepinsky (1950). Acta Cryst. 3, 160.

Farquhar and Lipsen (1946). Proc. Phys. Soc. 58, 200.

Furberg (1950). Acta Cryst. 3, 325.

Groth (1910) "Chemische Kristallographie" 3 (Leipzig).

Hagg and Laurent (1946). J. Sci. Inst. 23, 155.

Hirst and Macbeth (1926). J.C.S. 22.

Hudson (1926). J.A.C.S. 48, 1424; 48, 1434.

Lipson and Cochran (1953) "The Determination of Crystal Structures" (Bell, London).

MacDonald (1950) Thesis. University of Edinburgh.

MacDonald and Beevers (1950). Acta Cryst. 3, 394.

MacDonald and Beevers (1952). Acta Cryst. 5, 654.

MacGeachin (1954) Publs Dept. Cryst. Y Min. (Barcelona) 1, 145, 1954

Pitt (1948). Acta Cryst. 1, 168.

Reeves (1950). J.A.C.S. 72, 1499.

Reeves (1951) in Advances in Carbohydrate Chemistry 6.

Stern (1950), Thesis. University of Edinburgh.

Wilson (1942). Nature, 150, 152.

APPENDIX A.

h k l	F_o	Fc	h k l	F_o	F_c
1 0 0	25.5	34.5	$\bar{3}$ 0 2	18.1	17.8
2 0 0	18.6	-17.8	$\bar{2}$ 0 2	10.5	- 9.0
3 0 0	4.2	- 4.6	$\bar{1}$ 0 2	9.7	- 9.4
4 0 0	23.2	-22.2	0 0 2	32.9	34.3
5 0 0	4.9	4.6	1 0 2	30.2	-28.8
6 0 0	11.1	11.0	2 0 2	21.7	-21.2
7 0 0	5.1	- 6.1	3 0 2	8.7	- 8.9
8 0 0	2.5	- 2.5	4 0 2	17.4	-16.8
			5 0 2	1.1	0.7
$\bar{8}$ 0 1	6.7	- 6.3	6 0 2	0	- 2.4
$\bar{7}$ 0 1	9.3	9.6	7 0 2	16.3	-16.1
$\bar{6}$ 0 1	2.2	1.5	8 0 2	8.4	- 7.5
$\bar{5}$ 0 1	8.7	9.6			
$\bar{4}$ 0 1	4.8	- 4.2	$\bar{8}$ 0 3	2.6	- 3.1
$\bar{3}$ 0 1	47.1	-46.9	$\bar{7}$ 0 3	4.2	- 4.1
$\bar{2}$ 0 1	18.6	-18.1	$\bar{6}$ 0 3	14.8	-16.5
$\bar{1}$ 0 1	29.7	31.2	$\bar{5}$ 0 3	0	- 2.1
0 0 1	7.2	- 5.5	$\bar{4}$ 0 3	5.0	6.2
1 0 1	20.4	-20.2	$\bar{3}$ 0 3	11.9	-12.6
2 0 1	15.0	13.6	$\bar{2}$ 0 3	18.9	18.1
3 0 1	14.7	-13.7	$\bar{1}$ 0 3	50.3	48.5
4 0 1	11.2	11.2	0 0 3	6.9	- 7.1
5 0 1	10.5	11.2	1 0 3	13.5	-12.8
6 0 1	7.4	- 7.3	2 0 3	4.3	- 3.8
7 0 1	0	- 1.4	3 0 3	15.2	-15.2
8 0 1	6.6	- 6.7	4 0 3	22.6	-24.0
			5 0 3	24.5	23.6
$\bar{8}$ 0 2	6.6	- 6.9	6 0 3	0	- 1.0
$\bar{7}$ 0 2	3.2	- 4.7	7 0 3	3.6	3.5
$\bar{6}$ 0 2	14.2	15.1			
$\bar{5}$ 0 2	12.3	13.3	$\bar{7}$ 0 4	2.9	- 2.4
$\bar{4}$ 0 2	9.6	- 9.8	$\bar{6}$ 0 4	16.4	17.7

h k l	F_o	F_c	h k l	F_o	F_c
$\bar{5}$ 0 4	10.5	11.7	1 0 5	8.2	8.4
$\bar{4}$ 0 4	13.7	-13.8	2 0 5	1.9	1.9
$\bar{3}$ 0 4	12.9	12.9	3 0 5	15.9	-15.0
$\bar{2}$ 0 4	13.7	14.0	4 0 5	2.0	- 0.3
$\bar{1}$ 0 4	1.6	- 1.6	5 0 5	0	- 2.3
0 0 4	8.3	8.0			
1 0 4	9.3	- 8.7	$\bar{5}$ 0 6	5.5	6.4
2 0 4	13.0	-10.2	$\bar{4}$ 0 6	2.7	- 4.4
3 0 4	1.0	- 1.8	$\bar{3}$ 0 6	0	- 1.2
4 0 4	5.3	- 4.6	$\bar{2}$ 0 6	17.0	-17.1
5 0 4	6.1	- 4.9	$\bar{1}$ 0 6	18.3	-18.5
6 0 4	3.2	2.4	0 0 6	3.6	- 2.7
			1 0 6	7.0	- 6.8
$\bar{5}$ 0 5	7.8	- 7.6	2 0 6	6.2	- 5.0
$\bar{4}$ 0 5	0	- 0.1	3 0 6	7.9	9.0
$\bar{3}$ 0 5	0	- 3.0	4 0 6	7.7	8.5
$\bar{2}$ 0 5	0	0.2			
$\bar{1}$ 0 5	10.3	10.0	$\bar{1}$ 0 7	10.0	9.5
0 0 5	0	0.0	0 0 7	0	- 2.1
			1 0 7	2.9	- 1.7

h k l	$\|F_o\|$	A	B	$\|F_c\|$
0 2 0	41.2	32.3	27.2	42.2
1 2 0	39.6	32.9	-23.1	40.2
2 2 0	26.8	- 2.1	-27.7	27.8
3 2 0	1.4	- 0.7	2.6	2.9
4 2 0	9.7	- 3.1	- 8.7	9.2
5 2 0	15.3	-14.5	- 1.1	14.6
6 2 0	5.0	4.8	3.0	5.7
7 2 0	6.2	4.2	- 5.2	6.7
8 2 0	3.0	- 2.2	- 1.4	2.6
0 4 0	25.0	25.8	0.9	25.8
1 4 0	7.3	5.7	2.2	6.1
2 4 0	6.4	- 4.9	- 1.8	5.2
3 4 0	12.2	-11.6	1.4	11.8
4 4 0	13.2	-12.5	- 2.1	12.7
5 4 0	16.5	3.2	-16.0	16.3
6 4 0	7.8	6.1	- 2.6	6.6
7 4 0	0	- 0.2	1.4	1.4
0 6 0	22.7	7.7	21.1	22.4
1 6 0	3.8	2.7	3.1	4.1
2 6 0	0	- 1.8	1.4	2.3
3 6 0	5.4	- 2.8	4.2	5.0
4 6 0	13.9	- 7.2	-10.2	12.5
5 6 0	4.6	- 2.6	- 2.9	3.9
0 8 0	6.1	5.8	3.4	6.7
1 8 0	9.3	8.0	3.7	8.8
2 8 0	4.1	- 0.2	3.7	3.7
3 8 0	10.3	- 6.8	6.7	9.5

h k l	$\|F_o\|$	-A	B	$\|F_c\|$
1 1 0	26.1	5.0	26.2	26.7
2 1 0	44.9	-39.7	18.5	43.8
3 1 0	8.0	4.8	4.5	6.6
4 1 0	6.4	5.5	- 0.7	5.5
5 1 0	9.6	- 1.8	9.2	9.4
6 1 0	13.6	4.6	12.5	13.3
7 1 0	5.4	- 5.1	3.7	6.3
8 1 0	0	- 0.3	1.6	1.6
1 3 0	6.0	- 4.3	- 5.7	7.1
2 3 0	24.5	2.4	24.8	25.1
3 3 0	20.7	-20.7	- 3.8	21.0
4 3 0	15.3	- 7.8	-12.7	14.9
5 3 0	7.8	4.8	6.3	7.9
6 3 0	4.8	3.9	3.8	5.4
7 3 0	9.9	6.3	9.0	11.0
1 5 0	13.6	-14.1	0.9	14.1
2 5 0	2.8	- 3.5	- 0.3	3.5
3 5 0	9.7	3.3	8.8	9.4
4 5 0	6.0	4.1	- 3.5	5.4
5 5 0	8.1	8.6	- 1.7	8.8
6 5 0	4.8	0.4	5.2	5.2
1 7 0	8.6	- 3.8	8.5	9.3
2 7 0	15.9	- 2.9	15.9	16.2
3 7 0	2.2	1.2	0.9	1.5
4 7 0	4.3	- 0.4	- 3.9	3.9

APPENDIX B.

Values of $|F_{0kl}|$ observed by visual estimation, with F_{calc}.

h k l	$\|F_o\|$	A	B	$\|F_c\|$
0 2 1	25.4	-18.5	-12.8	22.5
0 2 2	8.2	- 0.4	6.4	6.4
0 2 3	10.9	5.9	-10.7	12.2
0 2 4	8.9	7.6	0.5	7.6
0 2 5	15.0	17.0	0.9	17.0
0 2 6	8.2	- 1.1	7.3	7.4
0 2 7	6.5	5.4	- 3.1	6.2
0 2 8	3.5	- 0.1	1.9	1.9
0 4 1	2.8	1.6	2.4	2.9
0 4 2	13.5	14.9	- 9.0	17.4
0 4 3	2.9	- 0.4	- 1.8	1.8
0 4 4	2.4	2.0	- 0.9	2.2
0 4 5	11.4	6.8	7.1	9.8
0 4 6	2.3	- 3.0	0.5	3.0
0 4 7	4.4	3.1	0.7	3.2
0 6 1	9.0	- 9.2	4.3	10.1
0 6 2	13.8	11.7	3.9	12.3
0 6 3	7.2	2.6	- 6.6	7.1
0 6 4	4.9	3.3	4.1	5.3
0 6 5	5.3	4.1	- 4.0	5.7
0 6 6	2.6	1.2	- 2.9	3.1
0 8 1	4.8	3.1	- 6.5	7.2
0 8 2	3.2	2.9	- 2.9	4.1
0 8 3	5.7	6.3	0.4	6.3
0 8 4	5.7	- 3.6	5.3	6.4
0 8 5	1.9	2.0	1.0	2.2

h k l	$\|F_o\|$	-A	B	$\|F_c\|$
0 1 1	25.7	- 3.5	27.5	27.7
0 1 2	22.0	- 7.9	19.7	21.2
0 1 3	15.1	- 7.5	-13.3	15.2
0 1 4	6.1	- 0.6	6.5	6.5
0 1 5	5.1	- 2.4	2.9	3.8
0 1 6	7.9	7.3	4.0	8.3
0 1 7	1.8	3.5	1.0	3.6
0 1 8	0	1.2	1.0	1.5
0 3 1	6.9	8.5	- 1.8	8.7
0 3 2	20.0	10.5	14.2	17.7
0 3 3	11.8	- 2.7	-10.4	10.7
0 3 4	4.2	- 4.1	- 1.7	4.4
0 3 5	9.9	- 9.3	2.1	9.5
0 3 6	2.1	1.0	- 0.8	1.3
0 3 7	2.2	+ 1.5	0.8	1.7
0 3 8	2.6	- 1.4	0.8	1.6
0 5 1	10.3	1.0	- 9.9	9.9
0 5 2	11.9	10.0	3.3	10.5
0 5 3	9.1	- 7.2	- 4.2	8.3
0 5 4	1.1	2.4	1.6	2.9
0 5 5	4.4	1.7	4.1	4.4
0 5 6	5.1	4.5	3.5	5.7
0 5 7	3.8	0.2	2.7	2.7
0 7 1	2.7	3.9	- 1.2	4.1
0 7 2	10.1	6.0	2.0	6.3
0 7 3	7.6	- 6.7	2.9	7.3
0 7 4	5.0	3.5	3.4	4.9
0 7 5	1.1	0.2	0.2	0.3
0 7 6	1.3	1.0	- 0.8	1.3

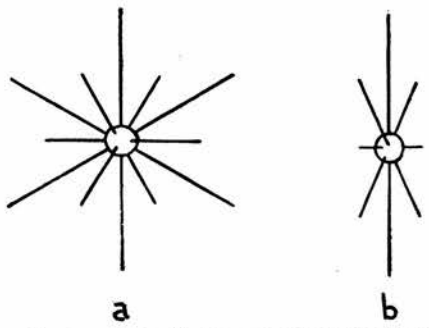


Fig. 1.— Vectores interatómicos para el anillo piranosa. Los vectores cortos en *a* están, alternativamente, por encima y por debajo del plano del papel.

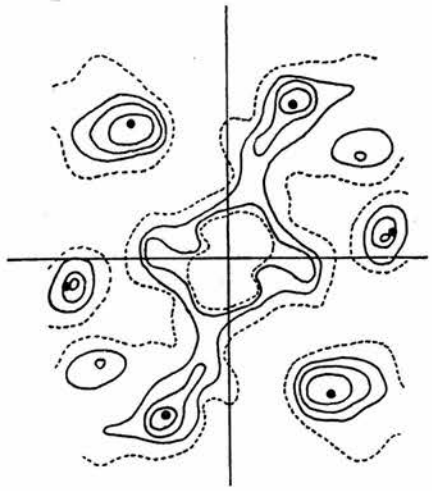


Fig. 2 — Sección de la función Patterson por el plano 7x — 8y— 2z 0. Están indicados los vectores 2'5 Å.

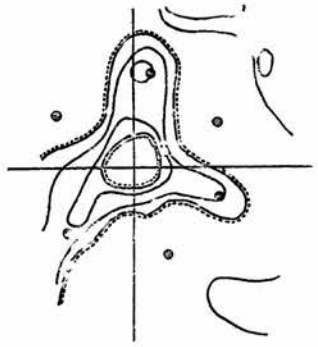


Fig. 3.— Sección del Patterson correspondiente a un plano paralelo al 7x — 8y — 2z 0 y distante 0'5 Å.

Figs.I, II and III. From Publs.Dept.Crist.Y Min. (Barcelona) 1, 145, 1954.

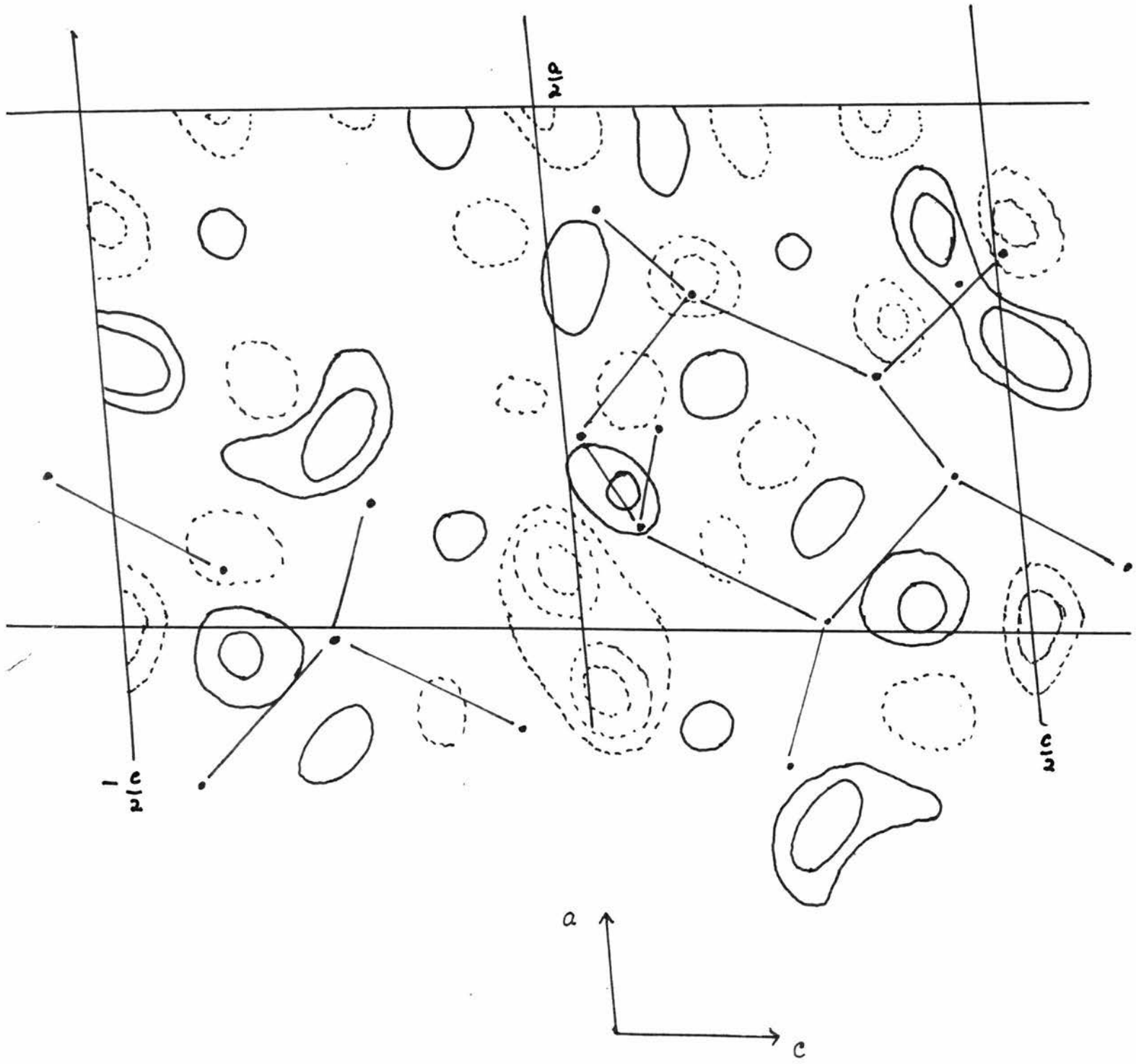


Fig.IV. (F_o-F_c) projection down the b-axis. The contour interval is $0.25eA^{-2}$. The zero contour is omitted. All atoms have been removed.

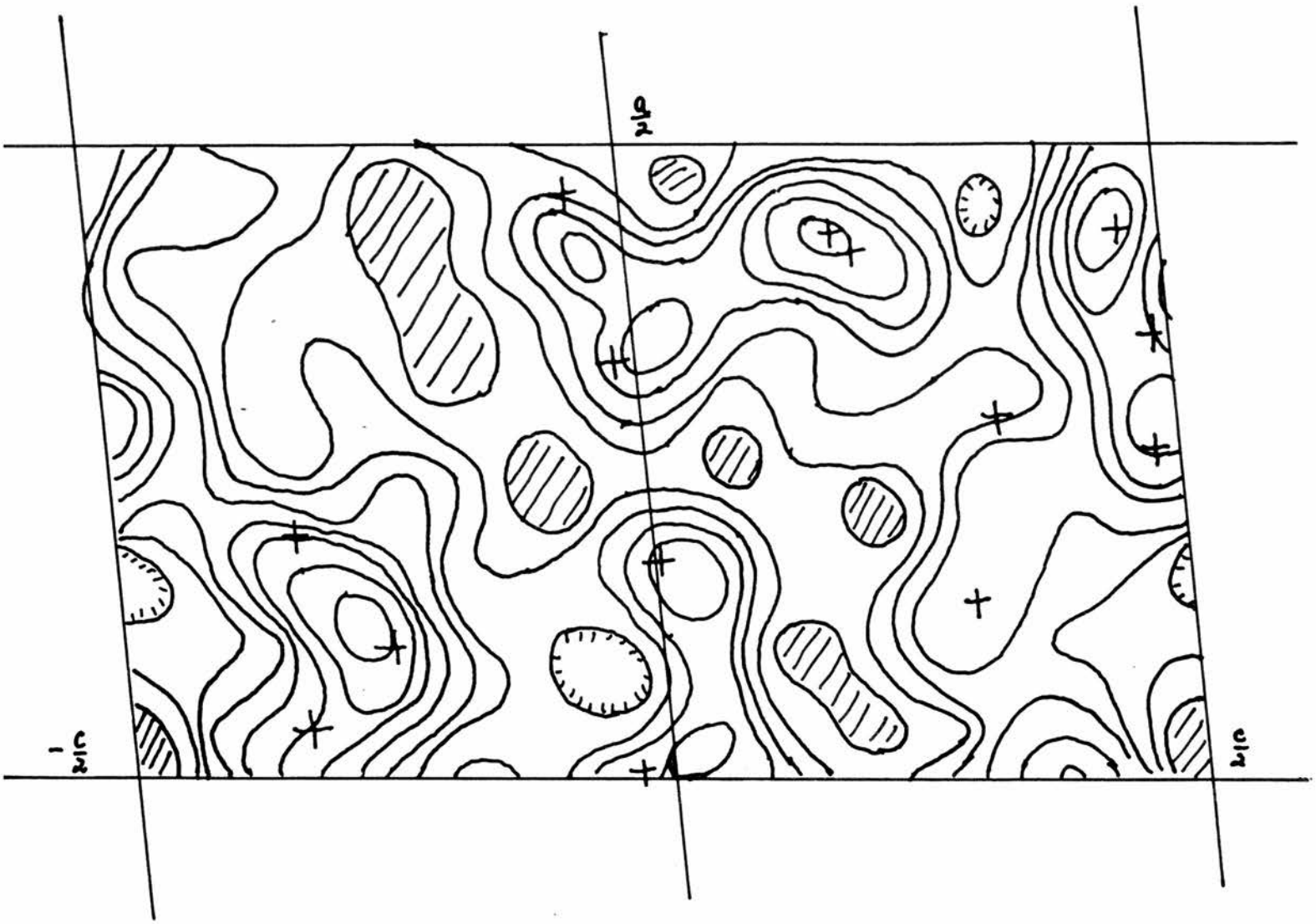


Fig.V. (F_o-F_c) projection down the b-axis. Contours as Fig.IV, but the zero contour not omitted. The hydrogen atoms are not removed, and their assumed positions are marked by crosses.

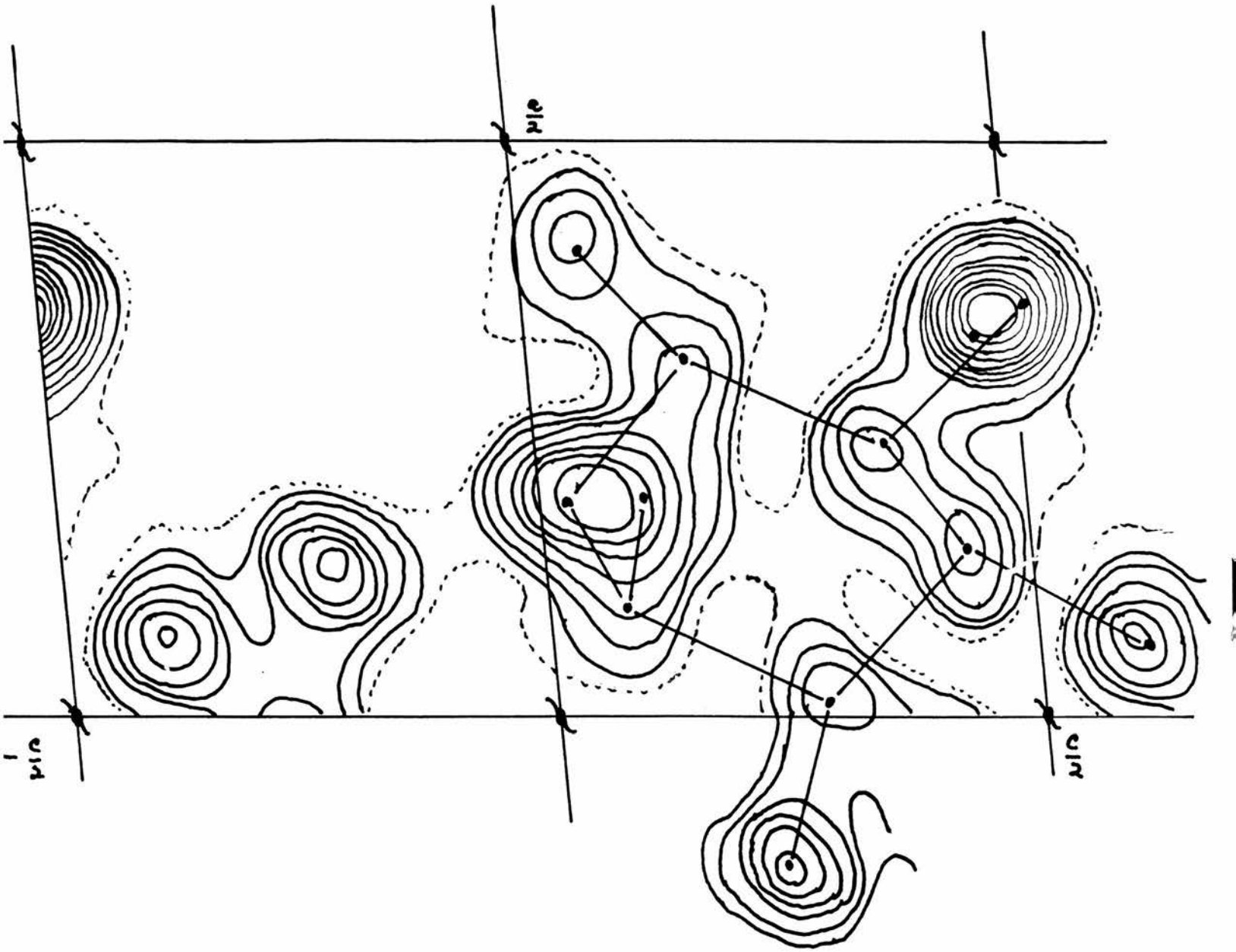


Fig.VI. Electron density projection down the b axis. Contours at 2 eA^{-2}. The 2 eA^{-2} contour is dotted.

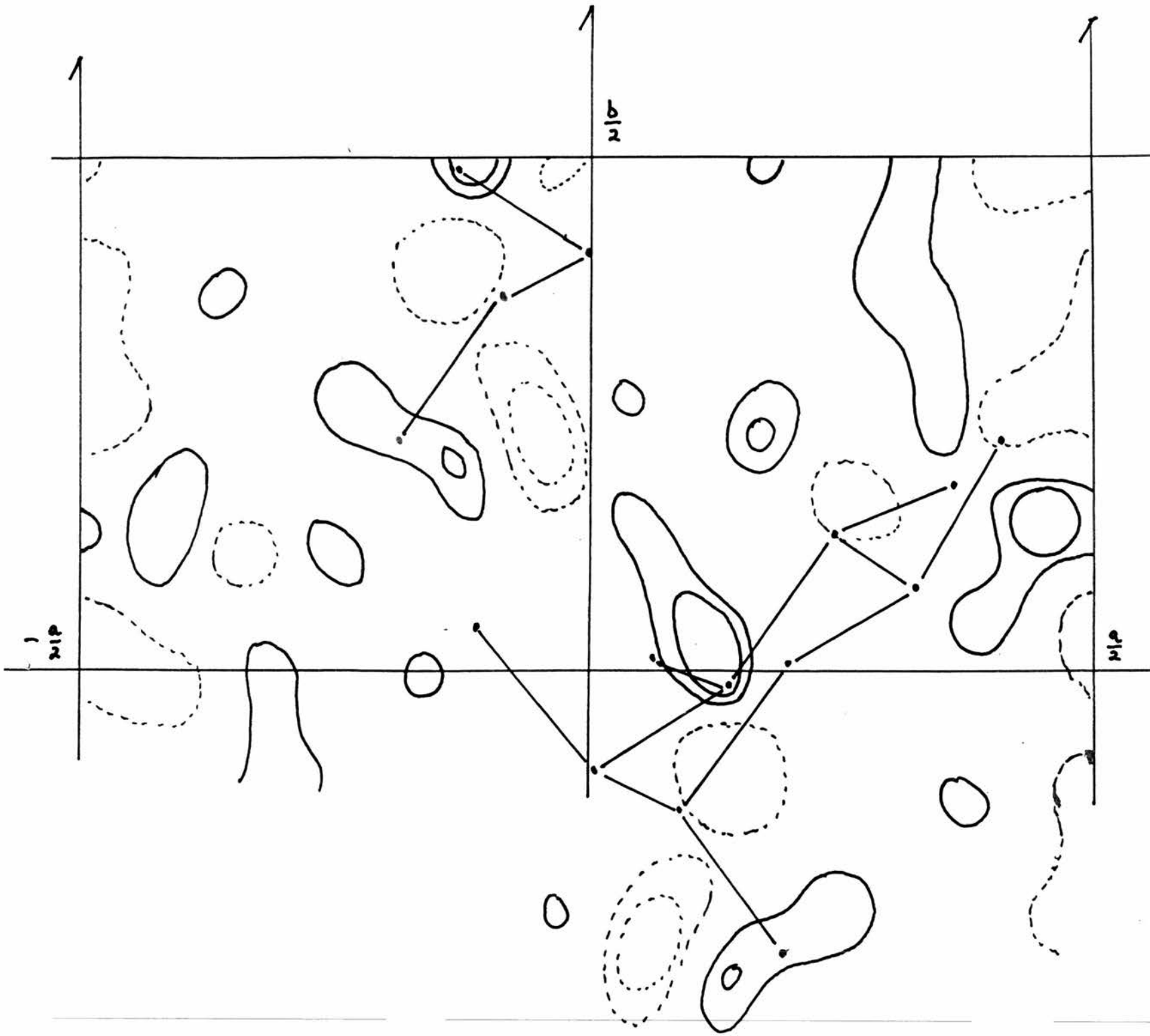


Fig.VII. (F_o-F_c) projection down the c axis. The contours are at 0.2 eA^{-2} and the zero contour is omitted.

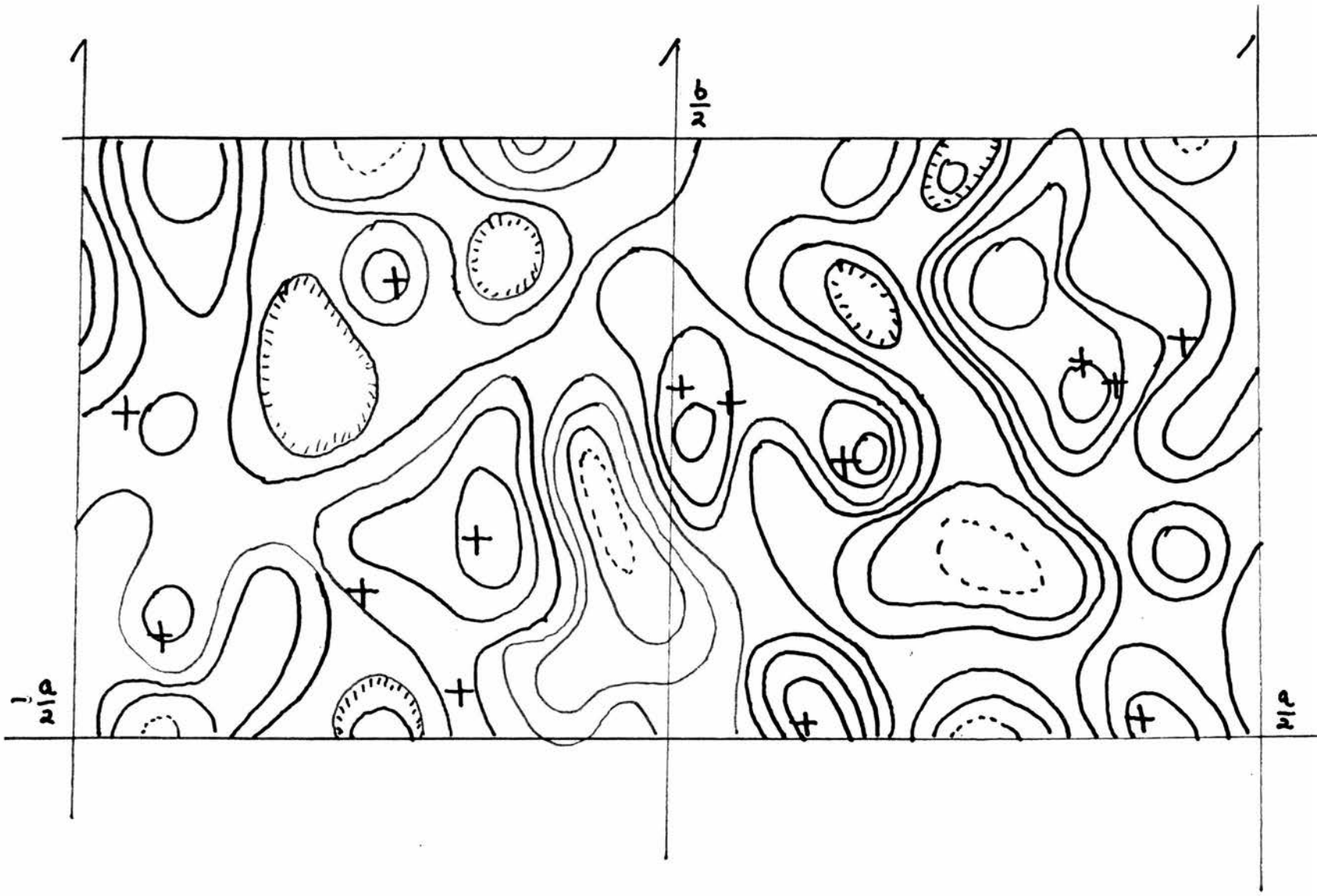


Fig.VIII. (F_o-F_c) projection down the c axis. The hydrogen atoms are removed and the phases are those of carbon and oxygen only. Contours at 0.2 eA^{-3}; the zero contour is not omitted.

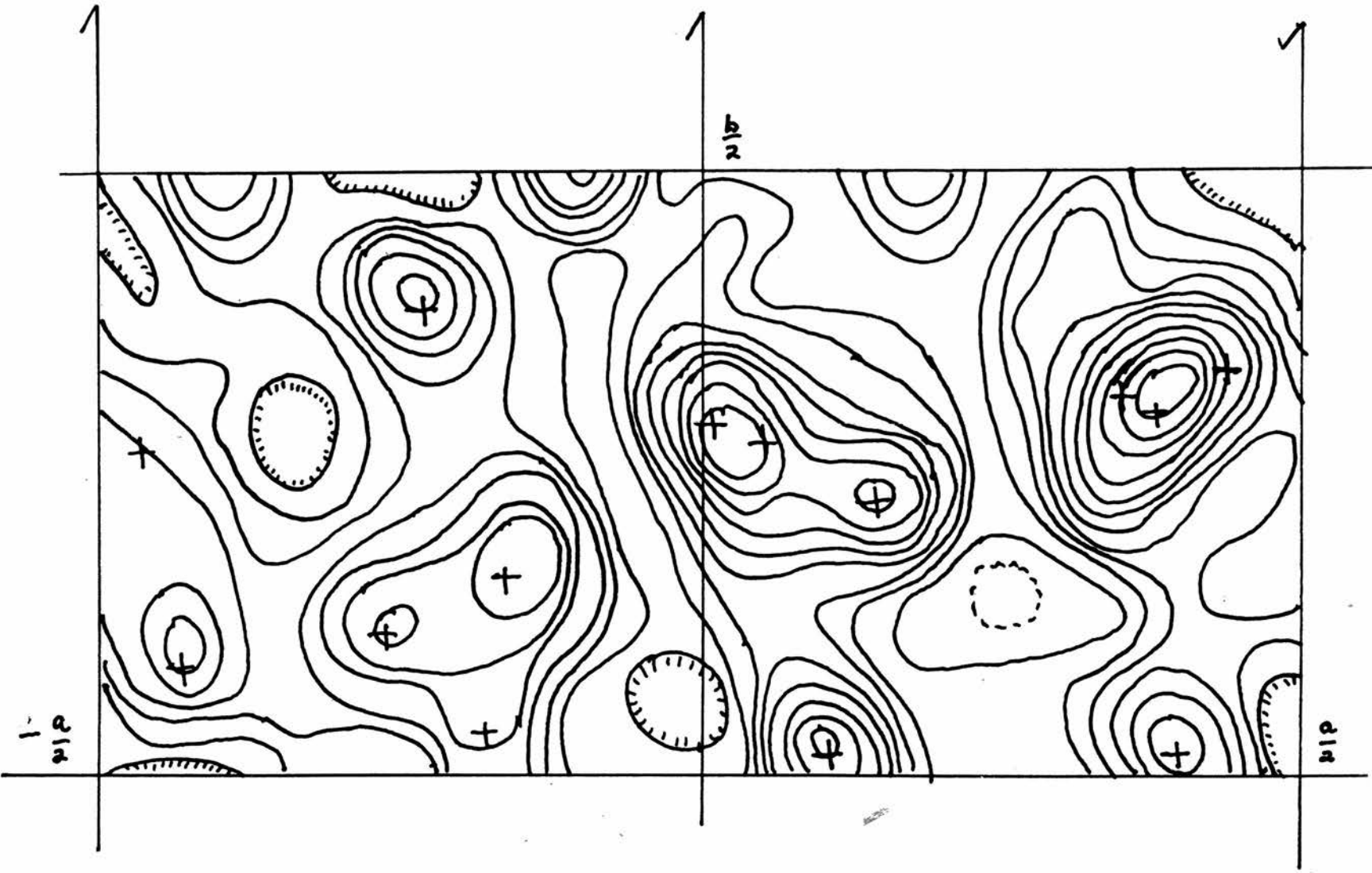


Fig.IX. (F_o-F_c) projection down the c axis. The hydrogen atoms are removed, but the phases used are those of the complete structure, including the hydrogen atoms in their assumed positions which are marked by crosses. Contours as Fig.VIII.

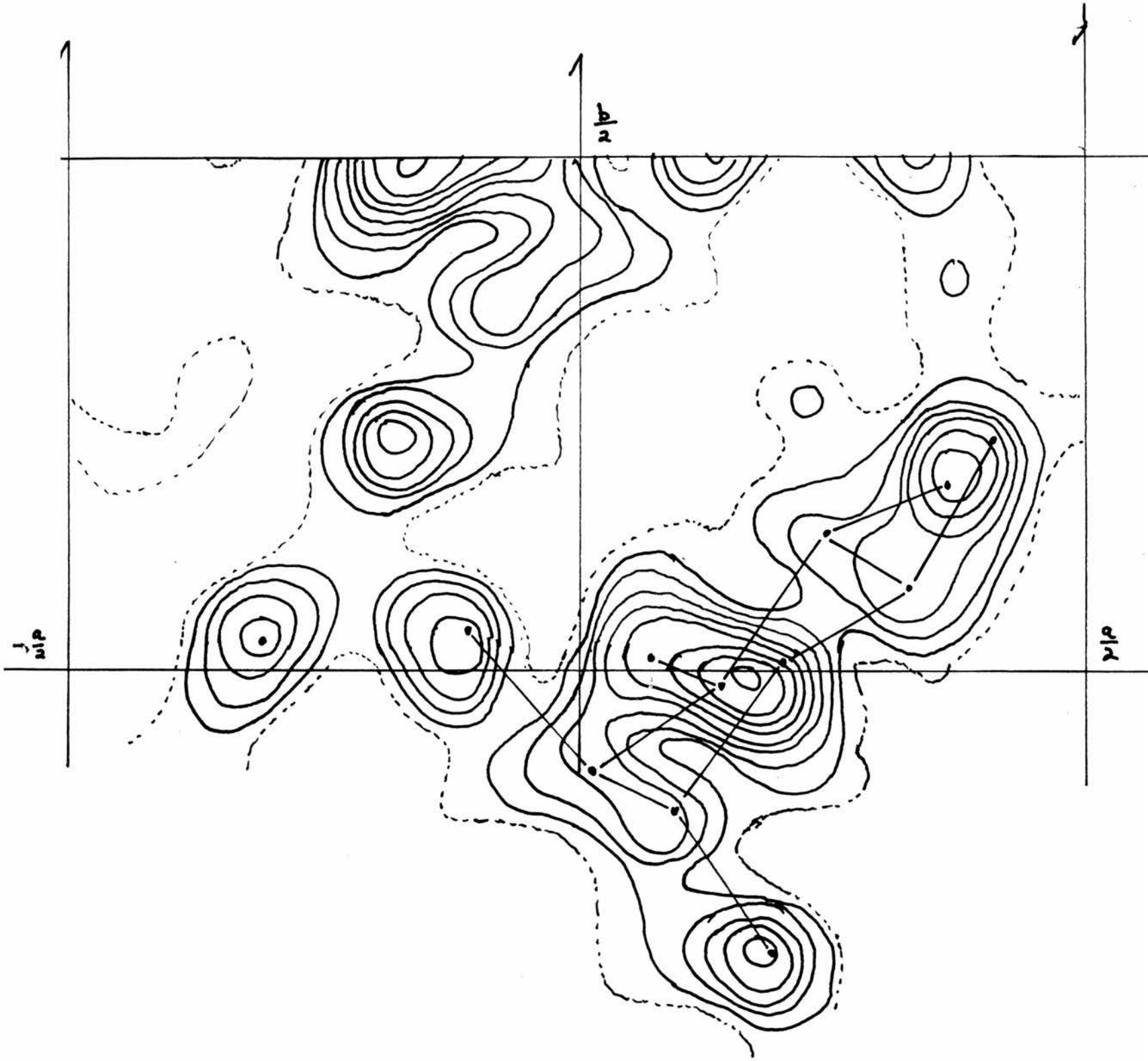


Fig.X. Electron density projection down the c axis. Contours at 2.0eA^{-2}, the 2.0eA^{-2} contour being dotted.

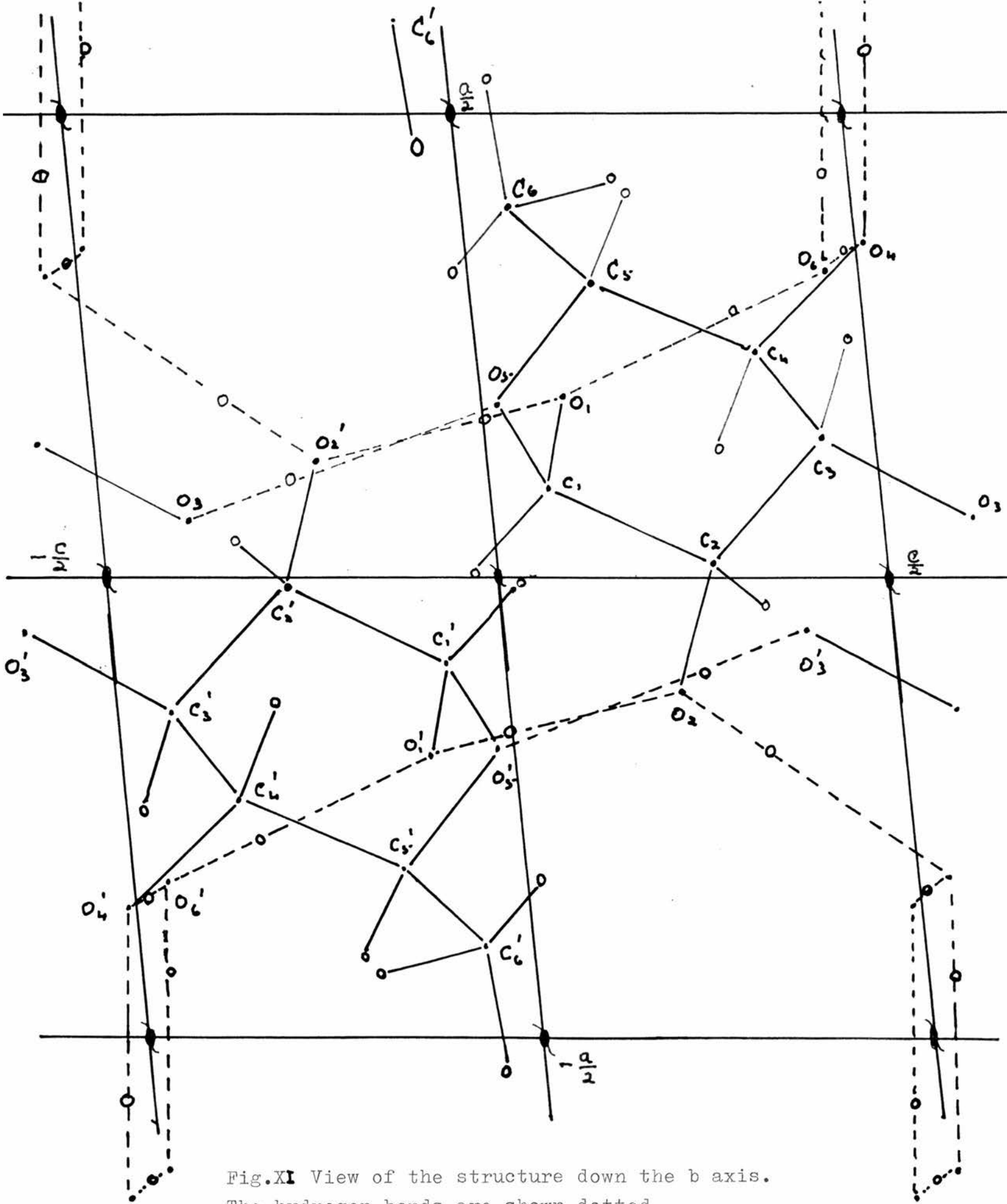


Fig.XI View of the structure down the b axis.
The hydrogen bonds are shown dotted.

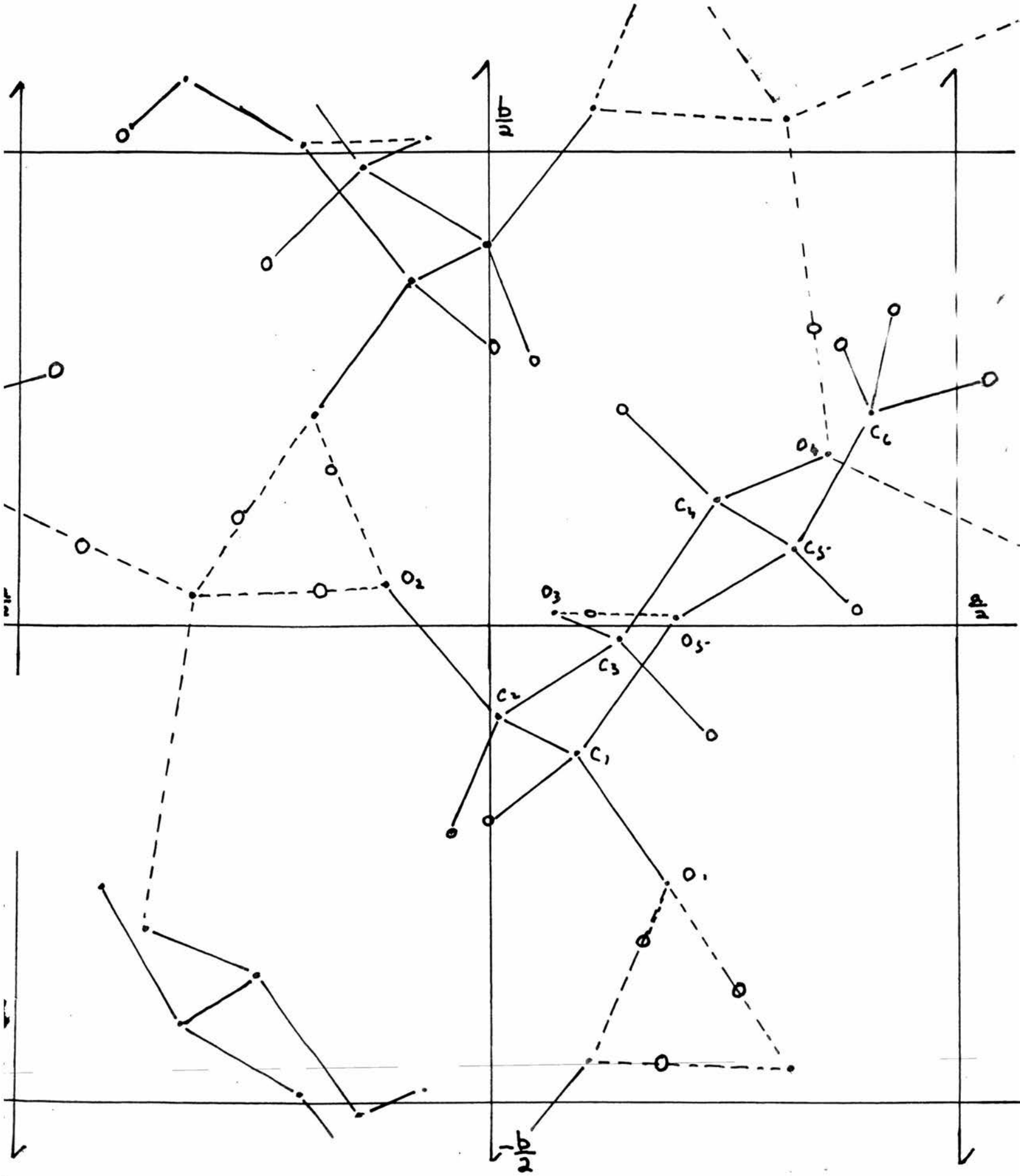


Fig.XII View of the structure down the c axis.